Multilevel Mediation Analysis:

Statistical Assumptions and Centering

by

Davood Tofighi

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2010 by the
Graduate Supervisory Committee:

Stephen G. West, Co-Chair
David P. MacKinnon, Co-Chair
Craig K. Enders
Roger E. Millsap

ARIZONA STATE UNIVERSITY

December 2010

ABSTRACT

Mediation analysis is a statistical approach that examines the effect of a treatment (e.g., prevention program) on an outcome (e.g., substance use) achieved by targeting and changing one or more intervening variables (e.g., peer drug use norms). The increased use of prevention intervention programs with outcomes measured at multiple time points following the intervention requires multilevel modeling techniques to account for clustering in the data. Estimating multilevel mediation models, in which all the variables are measured at individual level (Level 1), poses several challenges to researchers. The first challenge is to conceptualize a multilevel mediation model by clarifying the underlying statistical assumptions and implications of those assumptions on cluster-level (Level-2) covariance structure. A second challenge is that variables measured at Level 1 potentially contain both between- and within-cluster variation making interpretation of multilevel analysis difficult. As a result, multilevel mediation analyses may yield coefficient estimates that are composites of coefficient estimates at different levels if proper centering is not used. This dissertation addresses these two challenges. Study 1 discusses the concept of a correctly specified multilevel mediation model by examining the underlying statistical assumptions and implication of those assumptions on Level-2 covariance structure. Further, Study 1 presents analytical results showing algebraic relationships between the population parameters in a correctly specified multilevel mediation model. Study 2 extends previous work on centering in multilevel mediation analysis. First, different centering methods in multilevel analysis including centering within cluster with the cluster mean as a Level-2 predictor of intercept (CWC2) are discussed. Next, application of the CWC2 strategy to accommodate multilevel mediation models is explained. It is shown that the CWC2 centering strategy separates the between- and within-cluster mediated effects. Next, Study 2 discusses assumptions underlying a correctly specified CWC2 multilevel mediation model and defines between- and within-cluster mediated effects. In addition, analytical results for the algebraic relationships between the population parameters in a CWC2 multilevel mediation model are presented. Finally, Study 2 shows results of a simulation study conducted to verify derived algebraic relationships empirically.

DEDICATION

I am very grateful for all the support from the wonderful group faculty at the Arizona State University: Steve West and Leona Aiken, David MacKinnon, Craig Enders, and Roger Millsap.

My heartfelt appreciation goes to Steve West and Leona Aiken who have been there for me as mentors and advisors through ups and downs.

To our parents Allen and Ellen and especially to my mother whose unconditional love and support have been our strength and motivation through our life.

Finally, my special gratitude to my best friend and wonderful wife, Stephanie, whose unselfishness, constant support and unconditional love has made everything possible for me.

ACKNOWLEDGMENTS

The research was supported by Award Number F31DA027336 from the National Institute on Drug Abuse.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter 1

BACKGROUND

Mediation analysis is a statistical approach used to understand how a treatment produces an effect through intervening variables (mediators). In a prevention program, it is hypothesized that the prevention program changes mediators in order to reduce risk or enhance protective factors. Changes in these factors in turn are expected to decrease substance use and other negative health outcomes. A mediated effect associated with a mediator is defined as the effect of the prevention program on the outcome transmitted through the mediator. Prevention programs have been designed to decrease the association of participants with deviant peers (Coie et al., 1993), change participants' perception of drug use norms (Dishion & Skaggs, 2000; Dishion & Owen, 2002), and increase the participant's ability to refuse invitations by peers to use drugs (Ellickson, McCaffrey, Ghosh-Dastidar, & Longshore, 2003). Each of these intervening processes is expected to contribute to decreased drug use by program participants.

Conceptual work on preventive interventions has emphasized targeting risk and protective factors to decrease substance use (Epstein, Bang, & Botvin, 2007). Outcomes of preventive interventions can be seen as the result of two processes (Chen, 1990): First, the action theory (or program theory) describes the effect of a prevention program on mediators. Second, the conceptual theory (or psychosocial theory) describes how the mediators affect the outcomes. Current statistical approaches to mediation simultaneously model these two processes, partitioning the effect of the prevention program into a direct effect and mediated (indirect) effects that operate through the risk and protective factors.

The use of mediation analysis in basic and applied research has been increasing exponentially (as of 2010, Baron and Kenny, 1986, has over 19,000 citations, according to Google Scholar, October 18, 2010). Mediation models for analyzing single-level data (i.e., a data structure without clustering due to groups or repeated measures have received extensive development (e.g., MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002). However, the majority of drug prevention programs occur in group settings (clusters) such as schools or community groups (Botvin & Eng, 1980; Ellickson et al., 2003; Lynagh, Schofield, & Sanson-Fisher, 1997). A complexity of group settings is that the individuals within a cluster (e.g., classroom; community group) are more homogenous, thereby violating the statistical independence assumption. Individuals within a group may produce more homogeneous responses because of within-cluster interaction, influences from shared environmental factors, and similarity of group members. Clustered data sometime are referred to as *multilevel* data because the data collection may occur at multiple levels. For example, the data on individuals are referred to as Level-1 data whereas the data on clusters (groups) are referred to as Level-2 data. More sophisticated techniques such as multilevel analysis that account for clustering are needed (Hawkins et al., 2008) or clustering may lead to invalid results (Krull & MacKinnon, 2001, 1999; Raudenbush & Bryk, 2002).

Longitudinal trials in which mediators and outcomes are measured at multiple time points following intervention raise similar issues with respect to proper estimation (Hedeker & Gibbons, 1994). Observations from each person over time tend to be similar, violating the assumption of statistical independence. Multilevel and longitudinal models also permit researchers to address a wider variety of interesting substantive questions than traditional single-level models (Brook, Whiteman, Finch, & Cohen, 1995; Dooley, Prause, Ham-Rowbottom, & Emptage, 2006; Orlando, Ellickson, McCaffrey, & Longshore, 2005)

A multilevel mediation model refers to a mediation model in which at least one of the variables in the model (i.e., independent variable, mediator, and outcome variable) is measured at Level 1. The focus of this dissertation is on a two-level mediation model with one independent variable, one mediator, and one response variable, where all of the variables are measured at Level 1. Such a model is referred to 1-1-1 mediation model (Krull & MacKinnon, 1999, 2001) — a more detailed description of 1-1-1 mediation model will be provided later.

There are some issues in 1-1-1 mediation models that warrant more attention. The first issue is to conceptualize such models in terms of the underlying assumptions and to investigate the impact of those assumptions on the specification of covariances between Level-2 random coefficients (residuals). The covariances between Level-2 random coefficients (residuals) are referred to as a *Level-2 covariance structure*.

It appears that there is no consensus among researchers on how to *correctly* specify a Level-2 covariance structure in a 1-1-1 multilevel mediation model. For example, Krull and MacKinnon (1999, 2001) examined a 1-1-1 mediation model in which the intercepts, not the slopes, were allowed to vary randomly across clusters; random intercepts were not allowed to covary across clusters. Kenny, Korchmaros, and Bolger (2003), however, presented a 1-1-1 mediation model in which random intercepts were allowed to vary and certain random intercepts were allowed to *covary* (i.e., random slopes were correlated). Zhang, Zyphur, and Preacher (2009) presented a 1-1-1 mediation model in which random intercepts and random slopes were allowed to vary; however, neither the random intercepts nor the random slopes were permitted to covary.

More importantly, there seems to be little discussion in the literature on the link between Level-2 covariance structure and the assumptions underlying 1-1-1 mediation models. Making certain assumptions about a 1-1-1 mediation model has implications on the Level-2 covariance structure of the model and vice versa; different assumptions will lead to different Level-2 covariance structures, which in turn may have implications on the interpretation as well as the estimation of the coefficients and quantities of interest such as mediated effects.

The second issue is centering predictors in 1-1-1 mediation models and the impact of centering on interpretation of the mediated effects at individual (Level 1) and cluster (Level 2) levels. One difficulty of interpreting the results in a 1-1-1 mediation model is when the independent variable X, mediator M, and response variable Y contain information at Level 1 and Level 2 and those variables covary at each level. If a variable is measured at a lower level (Level 1), it is very likely that the variable contains variation at that level (Level 1) and the upper level

(Level 2). In addition, variables measured at Level 1 can be aggregated to have substantive interpretation at Level 2. For example, in classroom settings, a student perception of drug use norm (Level-1 measurement) can be aggregated at the classroom level (Level 2) to represent peer norm of drug use. The coefficients associated with the variable that contain variation at more than one level (e.g., both Level 1 and 2) are composites of the effects at those levels. Such coefficients are *uninterpretable* composites of the effects at different levels. A solution to this problem is centering.

The purpose of this dissertation is to address these two issues. This dissertation consists of two studies. Study 1 discusses the assumptions underlying a 1-1-1 mediation model and the impact of those assumptions on the specification of the Level-2 covariance structure. More specifically, Study 1 examines the assumptions and the Level-2 covariance structure for two *uncentered* 1-1-1 mediation models. The first one is referred to as a *correctly specified* uncentered 1-1-1 model, in which Level-2 random coefficients (residuals) are not correlated across regression equations. The second model is the *misspecified* uncentered 1-1-1 mediation model considered by Kenny et al. (2003). It will be shown that the uncentered multilevel mediation model discussed by Kenny et al. is *misspecified* because the Level-2 residuals in that model are correlated across regression equations.

It should be noted that the primary purpose of Study 1 is to discuss the results by Kenny et al. (2003). Therefore, Study 1 focuses on the *uncentered* multilevel mediation because Kenny et al. used that model in their study without centering the predictors. Further, using the uncentered multilevel mediation model in Study 1 provides a benchmark with which to compare the results of the centered 1-1-1 mediation model in Study 2. Study 2 discusses centering in a multilevel mediation model and how centering changes the interpretation of coefficients and mediated effects in a multilevel mediation model.

Single Mediator Model

In single-level randomized controlled trials with two groups (intervention vs. control), a single mediator model is defined as follows. An independent variable (e.g., $X = 1$, if a person participates in resistance skill program, otherwise 0) is hypothesized to change a mediator (e.g., M = drug refusal skill) which, in turn, changes a dependent variable (e.g., Y = frequency of drug use). Three equations used to assess quantities in the single mediator model are shown below (Baron & Kenny, 1986; MacKinnon, 2008).

$$Y_i = d_1 + c X_i + \varepsilon_{1i} \tag{1}$$

$$M_i = d_2 + a X_i + \varepsilon_{2i} \tag{2}$$

$$Y_i = d_3 + c' X_i + b M_i + \varepsilon_{3i} \tag{3}$$

where Y_i is the outcome variable measured on individual *i*, X_i is an indicator variable that represents whether the i^{th} person received the intervention (1= program; 0 = control), and M_i is the mediator. The coefficient c in Equation 1 represents the total effect of the prevention program on drug use. The coefficient c' in Equation 3 represents the direct effect of the prevention program on drug use, controlling for the participants' refusal skills. The direct effect captures the difference between treatment and control group adjusted for participants' refusal skills and indicates the part of the program effect not accounted for by the mediator. Coefficient *b* describes the effect of refusal skills on drug use controlling for the program effect. Coefficient *a* in Equation 2 represents the degree to which the intervention increased refusal skills relative to the control group. ε_{2i} and ε_{3i} are residuals that are have a bivariate normal distribution with zero correlation. This implies that the two error terms are independently and normally distributed across individuals, an assumption that may not hold when the individuals are clustered and the single-level model represented by Equations 2 - 3 are used for estimation. d_1, d_2, and d_3 are intercepts.

The magnitude of the effect of the prevention program on decreasing substance use mediated by the individuals' refusal skills is represented by the product of two coefficients, *a b*. The estimated mediated effect is shown by $\hat{a}\,\hat{b}$ where "$\hat{}$" sign denotes the estimator of each respective coefficient. Another equivalent measure of the mediated effect is $c - c'$.

Effect Size Measures and Quantities of Interest

Effect size measures provide a meaningful way of comparing the mediated effect across mediation studies regardless of sample sizes (MacKinnon, 2008). One of the most common measures of effect size to gauge the mediated effect is the ratio of the mediated effect to the total effect (Alwin & Hauser, 1975). There are three algebraically equivalent expressions for the ratio of the mediated effect to the total effect: $a\,b/(c' + a\,b)$, $a\,b/c$, and $1 - c'/c$. The ratio of mediated effect to the total effect gauges the proportion of the total effect of the independent variable (X) on the response variable (Y) transmitted through the mediator variable (M). For example, in a drug prevention program, if the proportion of mediated effect is .30, it indicates that 30% of the decrease in drug use was due to the increase in participants' refusal skills.

Another related effect size measure is the ratio of the mediated effect to the direct effect (ab/c') (Sobel, 1982). For example, if this measure is equal to 0.33, a researcher can conclude that the size of the mediated effect is one third of the size of the direct effect. Another useful measure to identify the surrogate (mediator) endpoint is the ratio of the total effect to the effect of X on M: c/a (Buyse & Molenberghs, 1998). Surrogate end points are the variables used instead of an ultimate outcome when the ultimate outcome has a low occurrence frequency or takes a long time to occur. For a variable (M) to be considered as a surrogate end point for the ultimate dependent variable (Y), the ratio of the total effect (c) to the effect of X on M (a) is expected to be close to one. That is, the magnitude of the relation between the surrogate endpoint M and the independent variable X should be equal to the magnitude of the relation between the dependent Y and the independent variable X.

Hypothesis Testing

Testing hypotheses in a single mediator model has received extensive attention (e.g., Baron & Kenny, 1986; MacKinnon, 2008; MacKinnon, Fritz, Williams, & Lockwood, 2007; MacKinnon et al., 2002; MacKinnon, Lockwood, & Williams, 2004; Preacher & Hayes, 2008). In classical statistics, the goal of hypothesis testing is to test if a parameter or a function of parameters is significantly different from zero. Researchers have recently emphasized using confidence intervals (CIs) as well as reporting p values for hypothesis testing (Harlow, Mulaik, & Steiger, 1997; Wilkinson and the Task Force on Statistical Inference, 1999). While classical hypothesis testing provides reject/not-reject decision for null hypothesis using test statistics, CIs also provide an interval that represents uncertainty in estimating the quantities of interest in a single mediator model. CIs can also be used in hypothesis testing.

There exist numerous procedures to test the mediated effect. One of the more commonly used tests for the mediated effect is the four-step test to establish mediation by Kenny and colleagues (Baron & Kenny, 1986). The four steps are as follows:

1. The effect of X on Y must be significant. That is, the parameter c in Equation 1 must be significant.
2. The effect of X on M, a, must be significant.
3. The effect of M on Y, b, controlling for X must be significant. This step is necessary to ensure that M also causes change in Y controlling for the direct effect of X on Y.
4. If the effect of X on Y controlling for M, c', is significant, we have partial mediation. If c' is not significant, complete mediation is achieved.

The second approach is to use asymptotic properties of maximum likelihood (ML) estimator of the mediated effect, $\hat{a}\hat{b}$, and form a z test statistic that is asymptotically normally distributed. In this approach,

$$z = \frac{\hat{a}\hat{b}}{SE(\hat{a}\hat{b})}$$

where $SE(\hat{a}\hat{b})$ is the standard error of $\hat{a}\hat{b}$. As the sample size increases for nonzero $a\, b$, the z statistic converges in distribution to the standard normal distribution. There are various methods to calculate $SE(\hat{a}\hat{b})$. Sobel (1982) proposed one of the most commonly used methods to calculate $SE(\hat{a}\hat{b})$. Sobel used the multivariate delta method to derive the approximate standard error of the indirect effect in structural equation models. Sobel's formula is:

$$SE(\hat{a}\hat{b}) = \sqrt{(a\,SE(\hat{b}))^2 + (b\,SE(\hat{a}))^2}$$

Another approach to test mediation is based on deriving the distribution of the product of two random normal variables (MacKinnon et al., 2007). The asymptotic Sobel's z test is based on the assumption that $\hat{a}\hat{b}$ has a normal distribution. However, this assumption does not hold for small to moderate sample sizes. In general, the product of two random normal variables is not normally distributed (Aroian, 1947; Lomnicki, 1967). A more accurate way of testing the mediated effect is to derive the underlying distribution of the $\hat{a}\hat{b}$. Craig (1936) derived the moments (e.g., skewness and kurtosis) and approximated the distribution of the product of random variables.

However, when either $\hat{a}$ or $\hat{b}$ has a mean of zero, the distribution of $\hat{a}\hat{b}$ is approximately proportional to the Bessel function of the second kind of zero order with a purely imaginative argument. The shape of the distribution is symmetric around the mean of zero. When neither $\hat{a}$ or $\hat{b}$ has a mean of zero and

$\hat{a}$ and $\hat{b}$ are independent, the central moments of $\hat{a}\,\hat{b}$ up to fourth order are as follows:

$$\mu_1 = \text{mean} = a\,b$$

$$\mu_2 = \text{variance} = \left(a\,SE(\hat{b})\right)^2 + \left(b\,SE(\hat{a})\right)^2 + \left(SE(\hat{a})\,SE(\hat{b})\right)^2$$

$$\mu_3 = \frac{6\,t_{\hat{a}}\,t_{\hat{b}}}{(t_{\hat{a}}^2 + t_{\hat{b}}^2 + 1)^{3/2}}$$

$$\mu_4 = \frac{6\left[2\left(t_{\hat{a}}^2 + t_{\hat{b}}^2\right) + 1\right]}{(t_{\hat{a}}^2 + t_{\hat{b}}^2 + 1)^2}$$

where $t_{\hat{a}} = \hat{a} / SE(\hat{a})$ and $t_{\hat{b}} = \hat{b} / SE(\hat{b})$. Craig (1936) showed that $\mu_3 \leq (2/3)\sqrt{3}$ and $\mu_4 \leq 6$. When both t_a and t_b are zero, kurtosis reaches its maximum at six.

Several simulation studies have investigated the performance of these tests. MacKinnon et al. (2002) compared several tests and CIs for the mediated effect. MacKinnon et al.'s study showed that Baron and Kenny's (1986) four-step approach is very conservative with low power unless the sample size is very large. For example, when the direct effect is zero ($c' = 0$), for small, medium, and large effect sizes for both *a* and *b*, the minimum sample sizes to detect .80 power were $N = 20{,}886$, $N = 397$, and $N = 92$, respectively. For Sobel's (1982) test, when both *a* and *b* were zero, the test was conservative in Type I error rate. When either *a* or *b*, but not both was zero, the test Type I error rate was close to the nominal value of .05. For Sobel's test, when the direct effect was zero ($c' = 0$), for small, medium, and large effect sizes for *a* and *b*, the minimum sample sizes to detect .80 power were 667, 90, and 42, respectively.

Multilevel Mediation

In my dissertation, I consider a multilevel mediation model where the predictor, mediator, and response variable are all measured at Level 1. This model is called the $1 \rightarrow 1 \rightarrow 1$ model (Krull & MacKinnon, 2001). However, for clarity I will use the notation 1-1-1 instead. The reason is that the arrows in notation $1 \rightarrow 1 \rightarrow 1$ may imply that the mediation exists *only* at Level 1 thereby causing confusion between the levels of measurement and the levels at which the mediated effects may occur. For example, if the variables contain only within-cluster variability and do not show between-cluster variation, the mediation occurs solely at Level 1. However, in most cases, Level-1 variables potentially contain both within- and between-cluster effects. As a result, we may have mediation occurring at either Level 1 or Level 2. I use the notation 1-1-1 without arrows to indicate that the predictor, mediator, and response variables are all measured at Level 1. Whether the mediation exists at Level 1 or Level 2 will be revealed in the analysis.

In a 1-1-1 multilevel mediation model, the three numbers indicate that predictor (X_{ij}), mediator (M_{ij}) and outcome variable (Y_{ij}), respectively, are all measured at Level 1 (e.g., individual level as opposed to cluster level). The subscripts i and j denote individual- and cluster-level subscripts, respectively. There are also other types of multilevel mediation models. For example, in the 2-

1-1 model, X_j is measured at Level 2 where M_{ij} and Y_{ij} are measured at Level. The focus of my dissertation is on 1-1-1 multilevel mediation models.

To illustrate, consider a two-level 1-1-1 multilevel model with a single mediator where students (Level 1) are nested within classrooms (Level 2). In classroom j, X_{ij} indicates student *i*'s perception of peer smoking (social norm), M_{ij} measures student *i*'s intention to smoke, and Y_{ij} is student *i*'s self-report of the number of cigarettes per day during last 30 days (behavior) (Guo et al., 2007). Note that the measurements occur at student level (Level 1). The statistical representation of a 1-1-1 multilevel mediation model parallels that of a single-level model in Equations 1- 3 except the coefficients can vary across classrooms. The student-level (Level-1) population model is shown below:

$$Y_{ij} = d_{1j} + c_j X_{ij} + \varepsilon_{1ij} \tag{4}$$

$$M_{ij} = d_{2j} + a_j X_{ij} + \varepsilon_{2ij} \tag{5}$$

$$Y_{ij} = d_{3j} + c'_j X_{ij} + b_j M_{ij} + \varepsilon_{3ij} \tag{6}$$

In addition, the classroom-level (Level-2) population model is as follows:

$$\begin{aligned} d_{1j} &= d_1 + u_{d_{1j}} \\ d_{2j} &= d_2 + u_{d_{2j}} \\ d_{3j} &= d_3 + u_{d_{3j}} \\ c_j &= c + u_{c_j} \\ a_j &= a + u_{a_j} \\ c'_j &= c' + u_{c'_j} \\ b_j &= b + u_{b_j} \end{aligned} \tag{7}$$

Combining the student-level (Level-1) and classroom-level (Level-2) equations in Equations 4-7, I arrive at the following equations known as the "mixed model" because it contains both "fixed" and "random" coefficients in each equation:

$$Y_{ij} = d_1 + c\,X_{ij} + u_{c_j}\,X_{ij} + u_{d_{1j}} + \varepsilon_{1ij} \tag{8}$$

$$M_{ij} = d_2 + a\,X_{ij} + u_{a_j}\,X_{ij} + u_{d_{2j}} + \varepsilon_{2ij} \tag{9}$$

$$Y_{ij} = d_3 + c'\,X_{ij} + b\,M_{ij} + u_{d_{3j}} + u_{c'_j}\,X_{ij} + u_{b_j}\,M_{ij} + \varepsilon_{3ij} \tag{10}$$

where i and j are individual and cluster level subscripts, respectively. X_{ij}, M_{ij}, Y_{ij} are the intervention (predictor), mediator, and outcome variables for the i^{th} person in cluster j. d_{1j}, d_{2j}, d_{3j} , c_j, a_j, b_j, and c'_j represent Level-1 random (e.g., student level) coefficients. d_1, d_2, d_3, a, b, c' and c are the "fixed" effect coefficients representing the population means. The interpretation of the fixed effects is similar to the interpretation of the parameters in a single-level mediation model (e.g., a, b, c' and c). For example, a quantifies the average effect of student's perception of peer smoking (X_{ij}) on student's intention to smoke (M_{ij}) across all students and classrooms. The residual terms $u_{d_{1j}}, u_{d_{2j}}, u_{d_{3j}}, u_{c_j}, u_{a_j}, u_{b_j},$ and $u_{c'_j}$ are "random effects" that capture the between-cluster variability of the random coefficients (e.g., a_j) around the population means (fixed effects). The single-level mediation model estimates population means, whereas the multilevel mediation estimates population means as well as between-cluster variability around the population means. The mediation effects can be examined at both individual and cluster levels.

Multilevel mediation analysis offers opportunities to test mediation at both individual and cluster levels while accounting for clustering of observations (Krull & MacKinnon, 1999, 2001). To obtain ML estimates for nonlinear functions of parameters such as the mediated effect or the proportion of total effect that is mediated, one can substitute the ML estimates in the expressions for the quantities of interest using the invariance property of ML. The central problem arises when one needs to compute the standard error for the ML estimators of the quantities that are nonlinear functions of coefficients. The standard errors are needed to develop test statistics or construct confidence intervals. As noted before, a general procedure to obtain approximate standard errors in classical statistics is to use the multivariate delta method.

Centering in Multilevel Modeling

Centering is an important issue in both single-level and multilevel regression analysis. Centering is defined as putting a predictor in deviation form so that the mean of the centered predictor is zero. Centering does not change the standard deviation of a predictor. Centering predictors makes the interpretation of the coefficients more meaningful (Aiken & West, 1991). For example, in OLS regression, centering changes the interpretation of the intercept.

In OLS regression, the intercept is the mean of the outcome variable when predictors take on the value of zero. If the zero point is outside of the range of the values of a predictor (e.g., the predictor ranges from one to five), the intercept value is not meaningful in terms of the actual data collected. The solution to this intercept problem is centering. Centering the predictor in OLS regression changes the interpretation of the intercept. The intercept is now meaningful. The intercept is the mean of the outcome variable when the centered predictor is zero, or equivalently, the predictor is at its mean level.

Centering in a 1-1-1 multilevel model is important because it makes interpretation of some of the coefficients more transparent. The problem stems from the fact that "many variables [at Level 1] can be conceptualized at more than one level, making the clear interpretation of some multilevel models difficult" (MacKinnon, 2008, p. 272). That is, the variable raw score (i.e., not centered or manipulated) potentially contains both individual- (Level-1) and cluster-level (Level-2) information. For example, the raw score for the predictor X_{ij} measured at Level 1 can be written as follows $X_{ij} = (X_{ij} - \bar{X}_j) + \bar{X}_j$, where $(X_{ij} - \bar{X}_j)$ is the centered within cluster (CWC) score and $\bar{X}_j$ is the cluster mean. $(X_{ij} - \bar{X}_j)$ contains only the within-cluster effect while $\bar{X}_j$ contains the between-cluster effect. The effects involving the raw score for the variable thus contain both individual level (within-cluster) and cluster level (between-cluster) components that may operate differently at each level. The coefficient estimates associated with raw scores represent a weighted average of both within-cluster and between-cluster effects. For example, for the effect of X_{ij} on M_{ij} , the following holds:

$$a = \eta^2 a_b + (1 - \eta^2) a_w ,$$

where η^2 is the proportion of the total variance of X_{ij} that is due to clustering, a_b and a_w are between-cluster and within-cluster effect of X_{ij} on M_{ij}, respectively (Kreft, de Leeuw, & Aiken, 1995). In other words, if we do not use proper centering, the estimate of a is an uninterpretable composite of two coefficients. If the researchers' intent were to estimate within-cluster effects and they do not center X_{ij}, the estimate of the within-cluster effect would be biased. The bias is as follows:

$$\begin{aligned} a - a_w &= \eta^2 a_b + (1-\eta^2) a_w - a_w \\ &= \eta^2 a_b - \eta^2 a_w \\ &= \eta^2 (a_b - a_w) \end{aligned}$$

The amount of bias is zero when between- and within-cluster effects are equal. Because researchers do not know beforehand if the effects are equal, they must use an appropriate centering strategy to avoid obtaining a biased estimate of between- and within-person effects.

As shown earlier, a Level-1 variable can be aggregated to the cluster level by computing the cluster mean (e.g., $\bar{X}_j$). The aggregated score (cluster mean) may conceptualize a cluster-specific construct that is different from the individual construct (MacKinnon, 2008, p. 272). MacKinnon provided an example of such variables: social norms. At the individual level, social norms provide individual measures of such norms, whereas the aggregated measure of social norms describes each group's social norms. Another example is a daily diary study of the effect of stress on alcohol use. At Level 1 (within-person), daily measures describe the daily fluctuation of stress and alcohol use. At Level 2 (between-person), the aggregated measures (person means) describe chronic stress and chronic alcohol use. Without a proper centering strategy, the between-cluster (person) and within-cluster (person) effects in each of the previous examples are confounded in a single estimate (e.g., a). By using an appropriate centering strategy one can estimate the between-person (e.g., a_b) and within-person (e.g., a_w) effects, separately.

There have been numerous articles discussing various aspects of centering in multilevel modeling (Enders & Tofighi, 2007; Kreft et al., 1995; Raudenbush, 1989a, 1989b). In multilevel modeling, one can center a predictor in two ways: centering within cluster or context (CWC) and centering at the grand mean (CGM). In CWC, the centered score is equal to the raw score minus the mean of the cluster to which the raw score belongs. The CGM score is obtained by subtracting the grand mean (mean of the raw score across all individuals and clusters) from each raw score.

To understand centering in multilevel analysis, I use a hypothetical example using simulated data. The simulated data are from a daily diary study of 100 individuals in which daily measures of stress and alcohol use are collected over 70 days. The measure of stress ranges between zero (no stress) to four (maximum stress). Alcohol use is the number of drinks per day for each individual. This data set has a multilevel structure where the daily measures of stress and alcohol use are nested (clustered) within individuals. Daily stress is a Level-1 predictor.

In this hypothetical study, a researcher is interested in the relationship between daily stress and daily alcohol use. Using the daily diary study, the researcher can pose two research questions. One research question is whether there is a relationship between the daily level of stress and daily alcohol use. This is a within-person research question. One can also ask a between-person research question. Using the very same daily diary data, one can characterize the chronic level of stress and chronic alcohol use by calculating average stress scores and average alcohol use over 70 days for each person. The between-person question is if there is a relationship between the chronic level of stress and chronic alcohol use. To answer the between- and within-person research questions in multilevel modeling, one has to use a centering strategy that separates the between- and within-person effects of stress on alcohol use.

This centering strategy is termed CWC2 (Kreft et al., 1995). CWC2 uses stress CWC scores as a Level-1 predictor and person means on stress as a Level-2 predictor of intercept in multilevel analysis. The CWC2 centering strategy separates the between- and within-person effects of stress on alcohol use.

To understand the importance of two steps of CWC2 centering strategy in multilevel modeling, let us look at within-person regression lines for one person, Person 2 in this example. For this person, the mean stress score is 2.0 and mean alcohol use is 2.6. In Figure 1, two within-person regression lines for Person 2 are shown. The regression lines capture the relationship between daily stress and alcohol use for Person 2. In the left panel of Figure 1, the horizontal axis shows the (uncentered) raw stress score and the vertical axis represents alcohol use. The intercept for the regression of alcohol use on stress scores for this person is 2.1. This means that the average alcohol use for Person 2 is 2.1 drinks per day when this person's stress level is 0. However, 0 is outside of the range of stress scores for Person 2, and therefore the estimate of the intercept is not meaningful. Now let us look at the graph in the right panel of Figure 1 where the horizontal axis represents the centered stress score obtained by subtracting the mean stress (i.e., 2.0) from each stress score for Person 2. This approach implies that I used CWC centering strategy (first step in CWC2 strategy) as I subtracted the person's mean from this person's own stress scores. After estimating the regression line using the centered score, the intercept estimate changes. The intercept estimate equals 2.1, which is exactly equal to the mean (chronic) alcohol use for Person 2. Note that when I used CWC daily stress as a predictor, the intercept becomes meaningful and equals the mean (chronic) alcohol use for that person. In addition, using CWC scores did not change the slope, which quantifies the within-person relationship between daily stress and alcohol use. In the first step of CWC2 centering strategy, using CWC stress scores makes the intercept equal to the chronic level of alcohol use for each person; the within-person relationship between stress and alcohol use remains unchanged.

Using CWC stress scores as a predictor in multilevel regression in the first step of CWC2 centering strategy, one can capture only the within-person relationship between stress and alcohol use— the between-person relationship between chronic level stress and alcohol use is *lost*. To capture both between- and within-person effects, one needs to follow the second step in the CWC2 centering strategy. That is, one needs to add person means on stress as a Level-2 predictor in the multilevel analysis. The Level-2 analysis in multilevel modeling captures the between-person relationship between the chronic level of stress and alcohol use. In the Level-2 regression, the outcome variable is the Level-1 intercept and the predictor is the stress mean for each person. Note that the Level-1 intercept is equal to the chronic (mean) level of alcohol use because I used CWC stress scores as a predictor in the first step of CWC2 centering strategy. The Level-2 regression slope quantifies the relationship between chronic stress (Level-2 predictor) and chronic alcohol use (Level-1 intercept).

In summary, to answer both the between- and within-person research questions in multilevel modeling, the analyst has to use a centering strategy that separates the between- and within-person effects. The centering strategy that separates the between- and within-person effects is termed CWC2 (Kreft et al., 1995). CWC2 refers to a strategy that uses CWC scores as a Level-1 predictor and person means as a Level-2 predictor of intercept in multilevel analysis.

Centering in Multilevel Mediation Modeling

As discussed in the previous section, centering is important in multilevel analysis. This is especially true when (at least) one of the variables X, M, and Y in a 1-1-1 model contains both within- and between-cluster variation. Several authors showed that between- and within-cluster coefficient estimates can differ substantially (Enders & Tofighi, 2007; Kreft et al., 1995).

In the context of multilevel mediation modeling, MacKinnon (2008, Chapter 9) discussed various centering strategies for a 2-1-1 model. Zhang et al.

(2009) also discussed centering strategies for 2-1-1 and 1-1-1 multilevel mediation models emphasizing that lack of proper centering can lead to confounded estimates of between- and within-cluster effects. Zhang et al. showed that the CWC2 centering strategy would lead to separate estimates between and within-cluster fixed effects in 1-1-1 model. In addition, Zhang et al. argued that in 2-1-1 model, the sample mediated effect can only exist at the between-cluster level because the independent variable *X* is measured at Level 2 thus contains only between-cluster variation. Preacher, Zyphur, and Zhang (2010) discussed a general approach to mediation model at Level 1 and Level 2 within a structural equation modeling framework. They used latent variable centering approach proposed by Lüdtke et al. (2008). Preacher et al.'s approach has one important limitation. In general, *MSE*s for the latent variable centering method were higher than the *MSE*s for the centering using the observed means (Ludtke et al., 2008).

Chapter 2

PURPOSE OF THIS DISSERTATION

My dissertation has two aims:

Study 1

Study 1 discusses various assumptions underlying an *uncentered* 1-1-1 mediation model and the consequences of those assumptions for the specification of Level-2 covariance structure. First, Study 1 discussed assumptions and the impact of those assumptions on the Level-2 covariances structure for a *correctly specified* uncentered 1-1-1 mediation model. In addition, Study 1 examines the assumptions underlying the *uncentered* 1-1-1 mediation model considered by Kenny et al. (2003) and the implications of those assumptions on the Level-2 covariance structure discussed in their study. It will be shown that the uncentered 1-1-1 mediation model discussed by Kenny et al. is *misspecified* because some of the Level-2 residuals are correlated. More specifically, in Kenny et al.'s model, there is a correlation between a_j and b_j. This specification implies that the associated Level-2 residuals, u_{a_j} and u_{b_j} are also correlated. Consequently, the uncentered 1-1-1 mediation model is *misspecified* because the Level-2 residuals are correlated across Equations 9 and 10. In other words, the Level-2 covariance structure considered by Kenny et al. (2003) and Bauer et al. (2006) is based on the assumption the uncentered 1-1-1 multilevel mediation model is misspecified.

It should be noted that the purpose of Study 1 is to discuss the results of Kenny et al. (2003). Therefore Study 1 focuses on the *uncentered* 1-1-1 mediation because Kenny et al. used that model in their study. Further, using an uncentered

1-1-1 mediation model provides a benchmark for the discussion of the results of applying the CWC2 centering strategy to an uncentered 1-1-1 mediation model in Study 2

After elucidating the assumptions underlying a correctly specified uncentered 1-1-1 mediation model, I present analytical results showing algebraic relationships between population parameters in that model and compare the results with the results of Kenny et al. (2003). For example, I show that $c = a\,b + c^{'}$ holds in correctly specified uncentered 1-1-1 models. For a misspecified 1-1-1 model, however, the following holds: $c = a\,b + c^{'} + \sigma_{a_j, b_j}$ (Kenny et al., 2003). To assess if the derived algebraic relationships between population parameters are empirically accurate, I report the results of a small-scale simulation study.

In summary, Study 1 answers the following questions through analytic work:

1. What are the assumptions underlying a correctly specified uncentered 1-1-1 mediation model? What is the Level-2 covariance structure for such a model?
2. Is the uncentered 1-1-1 mediation model considered by Kenny et al. (2003) and Bauer et al. (2006) is correctly specified? What are the assumptions underlying the model in Kenny et al.'s study that lead to correlation between random coefficients a_j and b_j in Equations 5 and 6?

3. Does the expression $c = ab + c'$ hold in a correctly specified uncentered 1-1-1 mediation model? How does this relationship differ from result discussed by Kenny et al (2003) for the misspecified 1-1-1 mediation model?

4. In a correctly specified 1-1-1mediation model, what are the relationships between the variances of d_{1j}, c_j and ε_{1ij} and covariance between d_{1j} and c_j in Equation 4 and variances and covariances of d_{2j}, a_j, ε_{2ij}, d_{3j}, c'_j, b_j, ε_{3ij} in Equations 5 and 6?

5. In a correctly specified 1-1-1 mediation model, are the derived relationships between the fixed and random effects empirically accurate as verified by a simulation study?

Study 2

Study 2 discusses the extension of the CWC2 centering strategy to an uncentered 1-1-1 mediation model thereby separating the between- and within-cluster effects. I term this model a *CWC2 1-1-1 mediation model*. Study 2 discusses statistical assumptions and the implications of such assumptions on the Level-2 covariance structure for a *correctly specified* CWC2 1-1-1 mediation model. One interesting effect of applying the CWC2 centering strategy is that it provides an opportunity to separate *between-cluster* and *within-cluster mediated effects*. Formal definitions for the between-cluster and within-cluster mediated effects will be presented below. In addition, the CWC2 centering strategy changes the expression $c = ab + c'$ as parameters in this formula are decomposed into between and within-cluster effects. Now, the question is what the formulas are for relationships between the between- (within-) cluster total effect, the between- (within-) cluster mediated effect and between-(within-) cluster direct effect in a CWC2 1-1-1 mediation model? Study 2 presents analytical formulas that show the relation $c = ab + c'$ holds for between- and within-cluster fixed effects, separately. In addition, I derive relationships between the random effects in a CWC2 1-1-1 mediation model. Finally, a simulation study will be conducted to assess the derived algebraic relationships for the fixed and random effects empirically.

In summary, Study 2 answers the following questions through analytic work:

1. How is the CWC2 centering strategy applied to an uncentered 1-1-1 mediation model? The resulting model is called CWC2 1-1-1 mediation model.
2. What are the statistical assumptions underlying a correctly specified CWC2 1-1-1 mediation model? What is the Level-2 covariance structure for such a model?
3. What is the definition of the between- and within-cluster mediated effects?
4. What are the relationships between fixed effects in a CWC2 1-1-1 mediation model? How does the formula $c = ab + c'$ change in a CWC2 1-1-1 mediation model?
5. What are the relationships between random effects in a CWC2 1-1-1 mediation model?

6. Are the derived algebraic relationships for the fixed and random effects empirically accurate as determined by a simulation study?

Chapter 3

STUDY 1

Statistical Assumptions in Correctly Specified Uncentered 1-1-1 Mediation Models

Statistical models are simplifications of the real world. As such, researchers make assumptions that are often “mathematically convenient” rather than realistic. The equations describing an uncentered 1-1-1 mediation model in Equations 4-7 are no exceptions. Failure to consider the underlying statistical assumptions and the implications of such assumptions in practice can potentially result in misleading statistical conclusions. This section discusses the assumptions underlying a *correctly specified* uncentered 1-1-1 mediation model and the implications of the assumptions on the Level-2 covariance structure for that model.

Before discussing the assumptions, I use a matrix representation of Equations 5 and 6 to show the concepts algebraically:

$$\begin{aligned} \text{Level-1:}\quad & \mathbf{M}_j = \mathbf{X}_{2j}\,\boldsymbol{\beta}_{2j} + \boldsymbol{\varepsilon}_{2j} \\ \text{Level-2:}\quad & \boldsymbol{\beta}_{2j} = \boldsymbol{\gamma}_2 + \mathbf{u}_{2j} \end{aligned} \qquad (11)$$

$$\begin{aligned} \text{Level-1:}\quad & \mathbf{Y}_j = \mathbf{X}_{3j}\,\boldsymbol{\beta}_{3j} + \boldsymbol{\varepsilon}_{3j} \\ \text{Level-2:}\quad & \boldsymbol{\beta}_{3j} = \boldsymbol{\gamma}_3 + \mathbf{u}_{3j} \end{aligned} \qquad (12)$$

where $\mathbf{M}_j$ and $\mathbf{Y}_j$ are $n_j \times 1$ vectors of the mediator and outcome variable for cluster j, respectively. $\mathbf{X}_{2j}$ is an $n_j \times 2$ matrix whose first column equals the constant one and the second column contains X_{ij} s. $\mathbf{X}_{3j}$ is an $n_j \times 3$ matrix whose first column contains all ones, the second column contains X_{ij} s, and the third column contains M_{ij} s. $\boldsymbol{\beta}_{2j} = \left(d_{2j} \;\; a_j\right)^T$ and $\boldsymbol{\beta}_{3j} = \left(d_{2j} \;\; c_j^{'} \;\; b_j\right)^T$ are Level-1 random coefficients. $\boldsymbol{\gamma}_2 = \left(d_2 \;\; a\right)^T$ and $\boldsymbol{\gamma}_3 = \left(d_3 \;\; c^{'} \;\; b\right)^T$ are Level-2 fixed coefficients. $\mathbf{u}_{2j} = (u_{d_{2j}} \;\; u_{a_j})^T$ and $\mathbf{u}_{3j} = (u_{d_{3j}} \;\; u_{c_j^{'}} \;\; u_{b_j})^T$ are Level-2 residuals.

The correct-specification assumption means that the uncentered 1-1-1 mediation model in Equations 5 and 6 (Equations 11 and 12) represents the "true" population causal model of the relationships between three variables, *X*, *M*, and *Y*. This assumption implies there is no omitted variable in the mediation model. In other words, given the 1-1-1 mediation model in Equations 11 and 12, the Level-1 and -2 residuals (errors) corresponding to the response variables $\mathbf{M}_j$ in Equations 11 and the ones associated with response variable $\mathbf{Y}_j$ in Equation 12 are independent from one another.

Statistically speaking, this means that for both Equations 5 and 6, ε_{2ij} and ε_{3ij} are identically and independently distributed (i.i.d.) across individuals and clusters. That is, within each equation, conditional on the predictors as well as Level-1 random coefficients, Level-1 residuals are i.i.d. More importantly, the residual terms are independent across Equations 5 and 6. More succinctly, we have:

$$\begin{pmatrix} \boldsymbol{\varepsilon}_{2j} \\ \boldsymbol{\varepsilon}_{3j} \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0}_{n_j \times 1} \\ \mathbf{0}_{n_j \times 1} \end{pmatrix}, \begin{pmatrix} \sigma_2^2 \, \mathbf{I}_{n_j} & \mathbf{0}_{n_j} \\ \mathbf{0}_{n_j} & \sigma_3^2 \, \mathbf{I}_{n_j} \end{pmatrix} \right) \tag{13}$$

where $\mathbf{0}_{n_j \times 1}$ is a zero column zero vector of size n_j, $\mathbf{0}_{n_j}$ is a zero square matrix of size n_j, and $\mathbf{I}_{n_j}$ is an identity square matrix of size n_j. Note that Equation 13 specifies the Level-2 covariance structure for the 1-1-1 mediation model. To emphasize, note that the covariance between the Level-1 residuals in Equations 11 and the ones in Equations 12 is zero.

In addition, the Level-2 residuals for a correctly specified uncentered 1-1-1 mediation model in Equations 11 and 12 are also assumed to be i.i.d with a multivariate normal distribution. That is,

$$\begin{pmatrix} \mathbf{u}_{2j} \\ \mathbf{u}_{3j} \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0}_{2\times1} \\ \mathbf{0}_{3\times1} \end{pmatrix}, \begin{pmatrix} \mathbf{\Sigma}_2 & \mathbf{\Sigma}_{2,3} \\ \mathbf{\Sigma}_{2,3} & \mathbf{\Sigma}_3 \end{pmatrix} \right) \tag{14}$$

$$\mathbf{\Sigma}_2 = \begin{pmatrix} \sigma^2_{d_{2j}} & \sigma_{d_{2j},a_j} \\ \sigma_{d_{2j},a_j} & \sigma^2_{a_j} \end{pmatrix}$$

$$\mathbf{\Sigma}_3 = \begin{pmatrix} \sigma^2_{d_{3j}} & \sigma_{d_{3j},c'_j} & \sigma_{d_{3j},b_j} \\ \sigma_{d_{3j},c'_j} & \sigma^2_{c'_j} & \sigma_{c'_j,b_j} \\ \sigma_{d_{3j},b_j} & \sigma_{c'_j,b_j} & \sigma^2_{b_j} \end{pmatrix}$$

$$\mathbf{\Sigma}_{2,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{15}$$

where $\mathbf{\Sigma}_2$ is a 2×2 *covariance matrix* of $\mathbf{u}_{2j}$ ($\boldsymbol{\beta}_{2j}$), $\mathbf{\Sigma}_3$ is a 3×3 covariance matrix of $\mathbf{u}_{3j}$ ($\boldsymbol{\beta}_{3j}$). $\mathbf{\Sigma}_{2,3}$ is a 2×3 covariance[1] matrix of $\mathbf{u}_{2j}$ ($\boldsymbol{\beta}_{2j}$) and $\mathbf{u}_{3j}$ ($\boldsymbol{\beta}_{3j}$) where the rows correspond to the elements in the random vector $\mathbf{u}_{2j}$ and the

[1] Anderson (2003, p. 35) uses the term covariance for $\mathbf{\Sigma}_3$ as well as $\mathbf{\Sigma}_{2,3}$, whether the matrices are symmetric or not. Similarly, I will use the term covariance to refer to the covariance matrix of a vector as well as the covariance matrix between two vectors.

columns correspond to the elements in the residual vector $\mathbf{u}_{3j}$. Note that the covariance matrix of $\boldsymbol{\beta}$ s and associated $\mathbf{u}$ s are equal. For example, $\sigma_{u_{d_{2j}},u_{a_j}}$ and σ_{d_{2j},a_j} are equal. Henceforth, I use these covariances and variances interchangeably throughout the text.

It should be noted that the covariance matrix in Equation 14 is the Level-2 covariance structure. Because the covariance matrix between $\mathbf{u}_{2j}$ and $\mathbf{u}_{3j}$ is zero (i.e., $\boldsymbol{\Sigma}_{2,3} = \mathbf{0}$), the Level-2 residuals $\mathbf{u}_{2j}$ and $\mathbf{u}_{3j}$ are *independent*. This implies that the 1-1-1 mediation model is correctly specified. Further, the Level-2 residuals have marginal multivariate normal distributions as follows (Anderson, 2003, Chapter 2):

$$\mathbf{u}_{2j} \sim N(\mathbf{0}_{2\times 1}, \boldsymbol{\Sigma}_2) \tag{16}$$

$$\mathbf{u}_{3j} \sim N(\mathbf{0}_{3\times 1}, \boldsymbol{\Sigma}_3) \tag{17}$$

Kenny et al.'s (2003) and Bauer et al.'s (2006) Approach

Kenny et al. (2003) and Bauer et al. (2006) presented a framework to describe the mediated effect in the uncentered 1-1-1 multilevel mediation model presented in Equations 4-7. Kenny et al. defined the mediated effect in terms of the product of two Level-1 random coefficients, a_j and b_j. Of importance is that Kenny et al. assumed that a_j and b_j are correlated, stating that this correlation is "a standard assumption within multilevel modelling" (p. 118). Using that assumption, Kenny et al. showed that

$$\mathrm{E}(a_j b_j) = ab + \sigma_{a_j,b_j} \tag{18}$$

where σ_{a_j,b_j} is the covariance between a_j and b_j.

In addition to lack of centering, the main issue in Kenny et al.'s (2003) result is that they assumed coefficients a_j and b_j are correlated when an uncentered 1-1-1 multilevel mediation is correctly specified. It should be noted that correlation between a_j and b_j implies that the associated Level-2 residuals, u_{a_j} and u_{b_j} are also correlated across regression equations corresponding to *M* and *Y*, respectively. As mentioned in the previous section, this means that the uncentered 1-1-1 mediation model considered by Kenny et al. and Bauer et al. (2006) is misspecified.

Relationship between Parameters in Correctly Specified Uncentered 1-1-1 Mediation Models

Having explicated the assumptions underlying a correctly specified 1-1-1 mediation model in Equations 5 and 6, I now discuss deriving algebraic relationships between fixed and random effects in (8) and the ones in (9) and (10). I intend to compare the relationship between fixed effects in a correctly specified 1-1-1 model with the relationship between the fixed effects in a misspecified 1-1-1 model. More specifically, this section shows the expression $c = a\,b + c'$ holds for a correctly specified model. This expression, however, does not hold a misspecified model (Kenny et al., 2003). The remainder of this section provides proofs for algebraic relationships between the fixed and random effects in an uncentered 1-1-1 mediation model.

My strategy is to make Equations 8 and 10 comparable (i.e., having the same predictors) by substituting the terms M_{ij} in (10) with the expressions that contain X_{ij}. In other words, I will reformulate Equation 10 so that it contains only X_{ij} as a predictor.

To reformulate Equation 10 so that it contains terms involving only X_{ij} s, I substitute Equation 9 into 10. To simplify the algebraic operations of the substitution, I first obtain the expressions for $b\,M_{ij}$ and $u_{bj}\,M_{ij}$ as follows:

$$b\,M_{ij} = b\,d_2 + a\,b\,X_{ij} + b\,u_{a_j} X_{ij} + b\,u_{i_2 j} + b\,\varepsilon_{2ij} \tag{19}$$

$$u_{b_j}\,M_{ij} = u_{b_j}\,d_2 + a\,X_{ij}\,u_{b_j} + u_{a_j} u_{b_j} X_{ij} + u_{b_j}\,u_{i_2 j} + u_{b_j}\,\varepsilon_{2ij} \tag{20}$$

Now, I substitute Equations 19 and 20 into Equation 10. Thus, the following holds:

$$\begin{aligned} Y_{ij} &= d_3 + c' X_{ij} + \left(b\,d_2 + a\,b\,X_{ij} + b\,u_{a_j} X_{ij} + b\,u_{i_2 j} + b\,\varepsilon_{2ij} \right) \\ &\quad + u_{i_3 j} + u_{c_j'} X_{ij} + \left(u_{b_j}\,d_2 + u_{b_j} u_{i_2 j} + a\,X_{ij}\,u_{b_j} + u_{a_j}\,u_{b_j} X_{ij} + u_{b_j}\,\varepsilon_{2ij} \right) + \varepsilon_{3ij} \\ &= \left(d_3 + b\,d_2 \right) + (c' + a\,b)\,X_{ij} + \left(u_{i_3 j} + b\,u_{i_2 j} + u_{b_j}\,d_2 + u_{b_j}\,u_{i_2 j} \right) \\ &\quad + (u_{c_j'} + a\,u_{b_j} + b\,u_{a_j} + u_{a_j}\,u_{b_j})\,X_{ij} + (\varepsilon_{3ij} + b\,\varepsilon_{2ij} + u_{b_j}\,\varepsilon_{2ij}) \end{aligned} \tag{21}$$

Note that Equation 21 reformulates Equation 10 so that it contains only X_{ij} as a predictor. Equation 8 and Equation 21 are now comparable.

In the next step, to derive the relationship between the fixed effects in Equations 8 and 21, I first obtain the expected values of both equations as follows:

$$\mathrm{E}(Y_{ij} \mid X_{ij}) = d_1 + c\,X_{ij} \tag{22}$$

$$\mathrm{E}(Y_{ij} \mid X_{ij}) = \left(d_3 + b\, d_2\right) + (c' + a\, b)\, X_{ij} \tag{23}$$

Because the terms on the left side of Equations 22 and 23 are equal, the ones on right side must also be equal. Thus, the corresponding terms from each equation (i.e., intercepts and slopes) must be equal for all X_{ij} or all intercepts and slopes must be equal to zero. Because the intercepts and slopes are not zero, the corresponding intercepts and slopes from each equation are equal. That is, the following holds:

$$\begin{aligned} d_1 &= d_3 + b\, d_2 \\ c &= c' + a\ b \end{aligned} \tag{24}$$

Interestingly, the above equation indicates that when the uncentered 1-1-1 multilevel mediation model is correctly specified (i.e., no correlated residuals), the equality $c = c' + a\, b$ holds as it does for a single-level mediation model. In addition, Equation 24 demonstrates the relationship between the intercept in Equation 4 and the intercepts in Equations 5 -6.

Additionally, I can obtain the similar algebraic expressions for the residuals terms $u_{i_1 j}$, u_{c_j}, and ε_{1ij}. To do so, I first calculate the residuals for Equations 8 and 21 , denoted by r_1 and r_1^*, respectively, using $Y_{ij} - \mathrm{E}(Y_{ij} \mid X_{ij})$. The residuals are as follow:

$$r_1 = u_{c_j} X_{ij} + u_{i_1 j} + \varepsilon_{1ij} \tag{25}$$

$$\begin{aligned} r_1^* = &\left(u_{i_3 j} + b\, u_{i_2 j} + u_{b_j}\, \gamma_{i_2 0} + u_{b_j}\, u_{i_2 j}\right) \\ &+ (u_{c_j'} + a\, u_{b_j} + b\, u_{a_j} + u_{a_j}\, u_{b_j})\, X_{ij} \\ &+ (\varepsilon_{3ij} + b\ \varepsilon_{2ij} + u_{b_j}\ \varepsilon_{2ij}) \end{aligned} \tag{26}$$

Note that r_1 and r_1^* contain both Level-1 and Level-2 residual terms. To obtain Level-2 residuals, I utilize the fact that Level-2 residual terms are fixed at Level 1 (i.e., within cluster). Therefore, by taking the expectation within a cluster, I can isolate Level-2 residual terms as follows:

$$\begin{aligned} u_1 &= \mathrm{E}(r_1 | u_{c_j}, u_{d_1 j}, X_{ij}) \\ &= \mathrm{E}(u_{c_j} X_{ij} + u_{d_1 j} + \varepsilon_{1ij} | u_{c_j}, u_{d_1 j}, X_{ij}) \\ &= u_{d_1 j} + u_{c_j} X_{ij} \end{aligned} \tag{27}$$

$$\begin{aligned} u_1^* &= \mathrm{E}(r_1^* | u_{d_{2j}}, u_{a_j}, u_{d_{3j}}, u_{c_j^{'}}, u_{b_j}, X_{ij}) \\ &= \mathrm{E}\big[(u_{d_{3j}} + b\, u_{d_{2j}} + u_{b_j} d_2 + u_{b_j} u_{d_{2j}}) \\ &\quad + (u_{c_j^{'}} + a\, u_{b_j} + b\, u_{a_j} + u_{a_j} u_{b_j}) X_{ij} + (\varepsilon_{3ij} + b\, \varepsilon_{2ij} + u_{b_j} \varepsilon_{2ij})\big] \\ &= (u_{d_3 j} + b\, u_{d_2 j} + u_{b_j} d_2 + u_{b_j} u_{d_2 j}) \\ &\quad + (u_{c_j^{'}} + a\, u_{b_j} + b\, u_{a_j} + u_{a_j} u_{b_j}) X_{ij} + \mathrm{E}(\varepsilon_{3ij}) + b\, \mathrm{E}(\varepsilon_{2ij}) + u_{b_j} \mathrm{E}(\varepsilon_{2ij}) \\ &= (u_{d_{3j}} + b\, u_{d_{2j}} + u_{b_j} d_2 + u_{b_j} u_{d_{2j}}) + (u_{c_j^{'}} + a\, u_{b_j} + b\, u_{a_j} + u_{a_j} u_{b_j}) X_{ij} \end{aligned} \tag{28}$$

where u_1 and u_1^* denote Level-2 residuals. As can be seen, the term associated with X_{ij} is the slope residual, whereas the remainder of the expression not associated with X_{ij} is the intercept residual. Because Equations 27 and 28 are equal, the corresponding residual terms in parentheses must be equal for all values of X_{ij} , or they all must be zero. Because the residual terms are not zero, the corresponding terms in each equation are equal. That is,

$$\begin{aligned} u_{d_{1j}} &= u_{d_{3j}} + b\, u_{d_{2j}} + u_{b_j} d_2 + u_{b_j} u_{d_{2j}} \\ u_{c_j} &= u_{c_j^{'}} + a\, u_{b_j} + b\, u_{a_j} + u_{a_j} u_{b_j} \end{aligned} \tag{29}$$

Finally, To obtain the relationship between Level-1 residuals, I first derive the expressions for the Level-1 residuals for Equations 8 and 21 by subtracting u_1 and u_1^* from r_1 and r_1^*, respectively. Let e_1 and e_1^* denote Level-1 residuals. Then,

$$e_1 = r_1 - u_1 = \varepsilon_{1ij}$$

$$e_1^* = r_1^* - u_1^* = \varepsilon_{3ij} + b\,\varepsilon_{2ij} + u_{b_j}\varepsilon_{2ij}$$

Thus, the following holds:

$$\varepsilon_{1ij} = \varepsilon_{3ij} + b\,\varepsilon_{2ij} + u_{b_j}\varepsilon_{2ij} \qquad (30)$$

Equations 29 and 30 show the relationships between Level-1 and Level-2 residual terms in Equation 4 and Level-1 and Level-2 residual terms in Equations 5-6.

Further, one can use expressions in (29) and (30) to obtain the variances for d_{1j} , c_j and ε_{1ij} and the covariance between d_{1j} and c_j in terms of the variances and covariances of the residual terms in Equations 5 and 6. Below is a detailed description of the analytic derivations for the variances of d_{1j} , c_j and ε_{1ij} and the covariance between d_{1j} and c_j .

First, I derive the variance for i_{1j} as follows:

$$\begin{aligned}
\sigma^2_{d_{1j}} &= Var(u_{d_{3j}} + b\,u_{d_{2j}} + u_{b_j} d_2 + u_{b_j} u_{d_{2j}}) \\
&= Var(u_{d_{3j}}) + b^2\,Var(u_{d_{2j}}) + d_2^2\,Var(u_{b_j}) + Var(u_{b_j}\,u_{d_{2j}}) \\
&+2\left\{ b\,Cov(u_{d_{3j}}, u_{d_{2j}}) + d_2\,Cov(u_{d_{3j}}, u_{b_j}) + Cov\big(u_{d_{3j}}, (u_{b_j}\,u_{d_{2j}})\big)\right\} \\
&= \sigma^2_{d_{3j}} + b^2\,\sigma^2_{d_{2j}} + d_2^2\,\sigma^2_{b_j} + \sigma^2_{b_j}\,\sigma^2_{d_{2j}} + 2\,d_2\,\sigma_{d_{3j},b_j}
\end{aligned}$$

The above equation is obtained by using the following relations:

$$
\begin{aligned}
Var(u_{b_j} u_{d_{2j}}) &= E(u_{b_j}\, u_{d_{2j}})^2 - \left(E(u_{b_j}\, u_{d_{2j}})\right)^2 \\
&= E(u_{b_j}^2\, u_{d_{2j}}^2) - \left(E(u_{b_j})\, E(u_{d_{2j}})\right)^2 \\
&= E(u_{b_j}^2)\, E(u_{d_{2j}}^2) = \sigma_{b_j}^2\, \sigma_{d_{2j}}^2
\end{aligned}
$$

Note that u_{b_j} and $u_{d_2 j}$ are independent (see Equation 15), and thus $u_{b_j}^2$ and $u_{d_2 j}^2$ are independent. As a result, $Cov\,(u_{b_j}, u_{d_2 j}) = 0$ and $Cov\,(u_{b_j}^2, u_{d_2 j}^2) = 0$. Further,

$$
\begin{aligned}
&Cov(u_{d_{3j}}, u_{d_{2j}}) = 0 \\
&Cov(u_{d_{3j}}, u_{b_j}) = \sigma_{d_{3j}, b_j} .
\end{aligned}
$$

Next, I derive $Cov\left(u_{d_{3j}}, (u_{b_j} u_{d_{2j}})\right)$, which is more elaborate:

$$
\begin{aligned}
Cov\left(u_{d_{3j}}, (u_{b_j} u_{d_{2j}})\right) &= E\,(u_{d_{3j}}\, u_{b_j}\, u_{d_{2j}}) - \left(E\,(u_{d_{3j}})\, E\,(u_{b_j}\, u_{d_{2j}})\right)^2 \\
&= E\,(u_{d_{3j}}\, u_{b_j} u_{d_{2j}}) + 0 \\
&= \int\!\!\int\!\!\int u_{d_{3j}}\, u_{b_j}\, u_{d_{2j}}\, f_{d_{3j}, b_j}(x, y) f_{d_{2j}}(z)\, dx\, dy\, dz \\
&= \int\!\!\int u_{d_{3j}}\, u_{b_j}\, f_{d_{3j}, b_j}(x, y)\, dx\, dy \int u_{d_{2j}}\, f_{d_{2j}}(z)\, dz \\
&= E\,(u_{d_{3j}}\, u_{b_j})\, E\,(u_{d_{2j}}) \\
&= \sigma_{d_{3j}, b_j} \times 0 = 0
\end{aligned}
$$

In the above derivation, I used the fact that $u_{d_2 j}$ is independent of both $u_{d_3 j}$ and u_{b_j} . Thus, it implies that it is independent of the product of $u_{d_3 j}$ and u_{b_j} . The variance of c_j is derived as follows:

$$
\begin{aligned}
\sigma_{c_j}^2 &= \mathrm{var}(u_{c_j'}) + \mathrm{var}(a\, u_{b_j}) + \mathrm{var}(b\, u_{a_j}) + \mathrm{var}(u_{a_j}\, u_{b_j}) \\
&+2\left\{a\ \mathrm{cov}(u_{c_j'}, u_{b_j}) + b\, \mathrm{cov}(u_{c_j'}, u_{a_j}) + \mathrm{cov}(u_{c_j'}, u_{a_j}\, u_{b_j})\right. \\
&\left.+\mathrm{cov}(a\, u_{b_j}, b\, u_{a_j}) + \mathrm{cov}(a\, u_{b_j}, u_{a_j}\, u_{b_j}) + \mathrm{cov}(b\, u_{a_j}, u_{a_j}\, u_{b_j})\right\} \\
&= \sigma_{c_j'}^2 + a^2 \sigma_{b_j}^2 + (b^2 + \sigma_{b_j}^2)\ \sigma_{a_j}^2 + 2\, a\, \sigma_{c_j', b_j}
\end{aligned}
$$

Note that in deriving the above equation, the following relations hold:

$$
\begin{aligned}
\operatorname{cov}(u_{c_j'}, u_{a_j} u_{b_j}) &= E(u_{c_j'} u_{a_j} u_{b_j}) \\
&= E(u_{a_j} u_{c_j'} u_{b_j}) \\
&= E(u_{a_j})\, E(u_{c_j'} u_{b_j}) \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{cov}(a u_{b_j}, b u_{a_j}) &= a\, b\ Cov\,(u_{b_j}, u_{a_j}) \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{cov}(a u_{b_j}, u_{a_j} u_{b_j}) &= a\ \operatorname{cov}(u_{b_j}, u_{a_j} u_{b_j}) \\
&= a\, E(u_{a_j} u_{b_j}^2) \\
&= a\, E(u_{a_j})\, E(u_{b_j}^2) \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{cov}(b\, u_{a_j}, u_{a_j} u_{b_j}) &= b \operatorname{cov}(u_{a_j}, u_{a_j} u_{b_j}) \\
&= b\, E(u_{a_j}^2\ u_{b_j}) \\
&= b\, E(u_{a_j}^2)\, E(u_{b_j}) \\
&= 0
\end{aligned}
$$

In addition, for Level-1 variance, σ_1^2, we have

$$
\begin{aligned}
\sigma_1^2 &= \operatorname{var}(\varepsilon_{1ij}) = \operatorname{var}(\, b\ \varepsilon_{2ij} + u_{b_j}\ \varepsilon_{2ij} + \varepsilon_{3ij}) \\
&= \operatorname{var}(b\ \varepsilon_{2ij}) + \operatorname{var}(u_{b_j}\ \varepsilon_{2ij}) + \operatorname{var}(\varepsilon_{3ij}) \\
&\quad +2\left\{ \operatorname{cov}(b\ \varepsilon_{2ij}, u_{b_j}\ \varepsilon_{2ij}) + \operatorname{cov}(b\ \varepsilon_{2ij}, \varepsilon_{3ij}) + \operatorname{cov}(u_{b_j}\ \varepsilon_{2ij}\ , \varepsilon_{3ij}) \right\} \\
&= b^2\ \mathrm{E}(\,\varepsilon_{2ij}^2) + \mathrm{E}(u_{b_j}^2 \varepsilon_{2ij}^2) + \mathrm{E}(\varepsilon_{3ij}^2) \\
&\quad +2\left\{\, b\ \mathrm{E}(u_{b_j}\ \varepsilon^2{}_{2ij}) + b\ \mathrm{E}(\,\varepsilon_{2ij}\ \varepsilon_{3ij}) + \mathrm{E}(u_{b_j}\ \varepsilon_{2ij}\ \varepsilon_{3ij}) \right\} \\
&= b^2 \sigma_2^2 + \sigma_{b_j}^2 \sigma_2^2 + \sigma_3^2 + 2\ \{\, 0 + 0 + 0\, \} \\
&= (b^2 + \sigma_{b_j}^2)\, \sigma_2^2 + \sigma_3^2
\end{aligned}
$$

Finally, one can derive the covariance between d_{1j} and c_j (i.e., $\sigma_{u_{1j}, u_{c_j}}$) as follows:

$$
\begin{aligned}
\sigma_{u_{i_1j},u_{c_j}} &= \mathrm{E}\Big[(u_{d_{3j}} + b\,u_{d_{2j}} + u_{b_j}\,d_2 + u_{b_j}u_{d_{2j}})\,(u_{c_j^{'}} + a\,u_{b_j} + b\,u_{a_j} + u_{a_j}u_{b_j})\Big] \\
&= \mathrm{E}\Big(u_{d_{3j}}\,(u_{c_j^{'}} + a\,u_{b_j} + b\,u_{a_j} + u_{a_j}\,u_{b_j}) + b\,u_{d_{2j}}(u_{c_j^{'}} + a\,u_{b_j} + b\,u_{a_j} + u_{a_j}u_{b_j}) \\
&\quad + u_{b_j}\,d_2(u_{c_j^{'}} + a\,u_{b_j} + b\,u_{a_j} + u_{a_j}u_{b_j}) + u_{b_j}u_{d_2j}(u_{c_j^{'}} + a\,u_{b_j} + b\,u_{a_j} + u_{a_j}u_{b_j})\Big) \\
&= \mathrm{E}\Big[u_{d_{3j}}\,u_{c_j^{'}} + a\,u_{d_{3j}}\,u_{b_j} + b\,u_{d_{3j}}\,u_{a_j} + u_{d_{3j}}\,u_{a_j}\,u_{b_j} + b\,u_{d_{2j}}\,u_{c_j^{'}} + a\,b\,u_{d_{2j}}\,u_{b_j} + b^2\,u_{d_{2j}}\,u_{a_j} \\
&\quad + b\,u_{d_{2j}}\,u_{a_j}\,u_{b_j} + d_2\,u_{c_j^{'}}\,u_{b_j} + a\,d_2\,u_{b_j}^2 + b\,d_2\,u_{a_j}u_{b_j} + d_2\,u_{a_j}u_{b_j}^2 \\
&\quad + u_{d_{2j}}\,u_{c_j^{'}}\,u_{b_j} + a\,u_{d_{2j}}\,u_{b_j}^2 + b\,u_{d_{2j}}\,u_{a_j}\,u_{b_j} + u_{d_{2j}}\,u_{a_j}\,u_{b_j}^2\Big] \\
&= \mathrm{E}[u_{d_{3j}}\,u_{c_j^{'}}] + a\,\mathrm{E}[u_{d_{3j}}\,u_{b_j}] + b\,\mathrm{E}[u_{d_{3j}}\,u_{a_j}] \\
&\quad + \mathrm{E}[u_{d_{3j}}\,u_{a_j}\,u_{b_j}] + b\,\mathrm{E}[u_{d_{2j}}\,u_{c_j^{'}}] + a\,b\,\mathrm{E}[u_{d_{2j}}\,u_{b_j}] \\
&\quad + b^2\,\mathrm{E}[u_{d_{2j}}\,u_{a_j}] + b\,\mathrm{E}[u_{d_{2j}}\,u_{a_j}\,u_{b_j}] + d_2\,\mathrm{E}[u_{c_j^{'}}\,u_{b_j}] \\
&\quad + a\,d_2\,\mathrm{E}[u_{b_j}^2] + b\,d_2\,\mathrm{E}[u_{a_j}\,u_{b_j}] + d_2\,\mathrm{E}[u_{a_j}\,u_{b_j}^2] \\
&\quad + \mathrm{E}[u_{d_{2j}}\,u_{c_j^{'}}\,u_{b_j}] + a\,\mathrm{E}[u_{d_{2j}}\,u_{b_j}^2] + b\,\mathrm{E}[u_{d_{2j}}\,u_{a_j}\,u_{b_j}] + \mathrm{E}[u_{d_{2j}}\,u_{a_j}\,u_{b_j}^2] \\
&= \sigma_{d_{3j},c_j^{'}} + a\,\sigma_{d_{3j},b_j} + 0 + 0 + 0 + 0 + b^2\,\sigma_{d_{2j},a_j} + 0 + d_2\,\sigma_{c_j^{'},b_j} \\
&\quad + a d_2\,\sigma_{b_j}^2 + 0 + 0 + 0 + 0 + 0 + \sigma_{d_{2j},a_j}\sigma_{b_j}^2 \\
&= \sigma_{d_{3j},c_j^{'}} + a\;\sigma_{d_{3j},b_j} + b^2\,\sigma_{d_{2j},a_j} + d_2\sigma_{c_j^{'},b_j} + a\;d_2\,\sigma_{b_j} + \sigma_{d_{2j},a_j}\sigma_{b_j}^2 \\
&= \sigma_{d_{2j},a_j}(b^2 + \sigma_{b_j}^2) + \sigma_{d_{3j},c_j^{'}} + a\,\sigma_{d_{3j},b_j} + d_2\,\sigma_{c_j^{'},b_j} + a\;d_2\,\sigma_{b_j}^2
\end{aligned}
$$

The above equation relies on the following equalities:

$$\mathrm{E}[u_{d_{3j}}\,u_{a_j}] = 0 \qquad (u_{d_{3j}} \text{ and } u_{a_j} \text{ are independent})$$

$$\mathrm{E}[u_{d_{2j}}\,u_{c_j^{'}}] = 0 \qquad (u_{d_{2j}} \text{ and } u_{c_j^{'}} \text{ are independent})$$

$$\mathrm{E}[u_{d_{2j}}\,u_{b_j}^2] = 0 \qquad (u_{d_{2j}} \text{ and } u_{b_j}^2 \text{ are independent})$$

$$
\begin{aligned}
\mathrm{E}[u_{d_{3j}}\,u_{a_j}\,u_{b_j}] &= \iiint u_{d_{3j}}\,u_{b_j}\,u_{a_j} f_{d_{3j},b_j,a_j}(x,y,z)\;dx\;dy\;dz \\
&= \iiint u_{d_{3j}}\,u_{b_j} f_{d_{3j},\,b_j}(x,y)\;u_{a_j} f_{a_j}(z)\;dx\;dy\;dz \\
&= \iint u_{d_{3j}}\,u_{b_j} f_{d_{3j},b_j}(x,y)\;dx\;dy \int u_{a_j} f_{a_j}(z)\,dz \\
&= \sigma_{d_{3j},b_j}\;E[u_{a_j}] = 0
\end{aligned}
$$

$$\mathrm{E}[u_{d_{2j}}\ u_{c_j'}] = 0 \qquad (u_{d_{2j}} \text{ and } u_{c_j'} \text{ are independent})$$

$$\mathrm{E}[u_{d_{2j}}\ u_{b_j}] = 0 \qquad (u_{d_{2j}} \text{ and } u_{b_j} \text{ are independent})$$

$$\begin{aligned}
\mathrm{E}[u_{d_2 j}\ u_{a_j}\ u_{b_j}] &= \iiint u_{d_2 j}\ u_{a_j}\ u_{b_j}\ f_{u_{d_2 j}\, u_{a_j}\, u_{b_j}}(x, y, z)\, dx\ dy\ dz \\
&= \iiint u_{d_2 j}\ u_{a_j}\ f_{u_{d_2 j}\, u_{a_j}}(x, y)\, dx\ dy\ \int u_{b_j} f_{u_{b_j}}(z)\, dz \\
&= \sigma_{d_{2j}, a_j}\ \mathrm{E}[u_{b_j}] \\
&= 0
\end{aligned}$$

$$\mathrm{E}[u_{b_j}^2] = \mathrm{var}(u_{b_j}) = \sigma_{b_j}^2$$

$$\mathrm{E}[u_{a_j} u_{b_j}] = \mathrm{E}[u_{a_j}]\,\mathrm{E}[u_{b_j}] = 0 \qquad (u_{a_j} \text{ and } u_{b_j} \text{ are independent})$$

$$\mathrm{E}[u_{a_j} u_{b_j}^2] = \mathrm{E}[u_{a_j}]\,\mathrm{E}[u_{b_j}^2] = 0 \qquad (u_{a_j} \text{ and } u_{b_j} \text{ are independent})$$

$$\begin{aligned}
\mathrm{E}[u_{d_{2j}}\ u_{c_j'}\ u_{b_j}] &= \int\!\int\!\int u_{d_{2j}}\ u_{c_j'}\ u_{b_j} f_{d_{2j}, c_j', b_j}(x, y, z)\ dx\ dy\ dz \\
&= \int\!\int\!\int u_{c_j'}\ u_{b_j} f_{c_j', b_j}(x, y)\ u_{d_{2j}} f_{u_{d_{2j}}}(z)\ dx\ dy\ dz \\
&= \int\!\int u_{c_j'}\ u_{b_j} f_{c_j', b_j}(x, y)\ dx\ dy\ \int u_{d_{2j}} f_{u_{d_{2j}}}(z)\ dz \\
&= \sigma_{c_j', b_j}\ \mathrm{E}[u_{d_{2j}}] \\
&= 0
\end{aligned}$$

$$\mathrm{E}[u_{d_{2j}}\ u_{b_j}^2] = \mathrm{E}[u_{d_{2j}}]\,\mathrm{E}[u_{b_j}^2] = 0 \qquad (u_{d_{2j}} \text{ and } u_{b_j} \text{ are independent})$$

$$\begin{aligned}
\mathrm{E}[u_{d_{2j}}\ u_{a_j}\ u_{b_j}^2] &= \int\!\int\!\int u_{d_{2j}}\ u_{a_j}\ u_{b_j}^2\ f_{u_{d_{2j}}, a_j, b_j}(x, y, z)\ dx\,dy\,dz \\
&= \int\!\int\!\int u_{d_{2j}}\ u_{a_j}\ f_{u_{d_{2j}}, a_j}(x, y)\ u_{b_j}^2\ f_{b_j}(z)\,dx\,dy\,dz \\
&= \int\!\int u_{d_{2j}}\ u_{a_j}\ f_{u_{d_{2j}}, a_j}(x, y)\ dx\,dy \int u_{b_j}^2\ f_{b_j}(z)\ dz \\
&= \sigma_{d_{2j}, a_j}\ \mathrm{E}[u_{b_j}^2] \\
&= \sigma_{d_{2j}, a_j} \sigma_{b_j}^2
\end{aligned}$$

Below is the summary of the derived relationships between the variances and covariances Level-1 and Level-2 residuals:

$$
\begin{aligned}
\sigma^2_{d_{1j}} &= b^2\,\sigma^2_{d_{2j}} + d_2^2\,\sigma^2_{b_j} + \sigma^2_{b_j}\sigma^2_{d_{2j}} + 2\,d_2\,\sigma_{d_{3j},b_j} \\
\sigma^2_{c_j} &= \sigma^2_{c'_j} + a^2\sigma^2_{b_j} + (b^2 + \sigma^2_{b_j})\,\sigma^2_{a_j} + 2a\,\sigma_{c'_j,b_j} \\
\sigma^2_1 &= (b^2 + \sigma^2_{b_j})\,\sigma^2_2 + \sigma^2_3 \\
\sigma_{d_{1j},c_j} &= \sigma_{d_{2j},a_j}(b^2 + \sigma^2_{b_j}) + \sigma_{d_{3j},c'_j} + a\;\sigma_{d_{3j},b_j} + d_2\,\sigma_{c'_j,b_j} + a\;d_2\,\sigma^2_{b_j}
\end{aligned}
\qquad (31)
$$

Simulation Study

In this section I describe a small-scale simulation study to assess whether the derived relationships between fixed effects and random effects in Equations 24 and 31 are empirically accurate. In the simulation study, I used Equations 24 and 31 to calculate the population values for the parameters d_1, c, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, and σ^2_1. Note that these parameters are not free parameters. They are a function of other parameters as shown in Equations 24 and 31. To emphasize, these parameters are referred to as *inferred* parameters as indicated in Kenny et al. (2003).

Then I use these population values to estimate bias, relative bias, and mean squared error (*MSE*) of the restricted maximum likelihood (REML) estimates of the parameters in Equation 8. For example, for the inferred parameter d_1, the estimate $(\hat{d}_1)$, the bias, relative bias, and *MSE* are obtained as follows:

$$\hat{d}_1 = \frac{\sum \hat{d}_1^{rep}}{n}$$

$$bias = \hat{d}_1 - d_1$$

$$relative\ bias = \frac{\hat{d}_1 - d_1}{d_1}$$

$$MSE = \frac{\sum(\hat{d}_1^{rep} - d_1)^2}{n-1}$$

where $\hat{d}_1^{rep}$ is the REML estimate for d_1 obtained from each simulation replication *rep,* and *n* is total the number of replications.

The purpose of this simulation study is to evaluate the accuracy of the derivation of the inferred population values d_1, c, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, and σ^2_1, not the accuracy of REML estimator of the parameters. The strategy used in this simulation study can be thought in terms of "reverse engineering" of the typical simulation design—the estimators' values are used to verify the accuracy of the associated population values. Note that this is based on the fact that the REML estimators for fixed effects and σ^2_1 are unbiased. In addition, the REML estimators for $\sigma^2_{d_{1j}}$ and $\sigma^2_{c_j}$ are asymptotically unbiased. Therefore, if the derived formulas for inferred parameters are accurate, then bias and relative bias for the inferred parameters in the simulation study will get close to zero as the number of clusters becomes "large".

Simulation design. The population model for the simulation study is based on the multilevel mediation model in Equations 9 and 10. The population parameters for the simulation are as follows: $c = 0.4725$, $a = 0.35$, $c' = 0.35$, $b = 0.35$, $\sigma^2_1 = 2.1225$, $\sigma^2_2 = 1$, $\sigma^2_3 = 1$, $\sigma^2_{d_{1j}} = 2.1225$, $\sigma^2_{d_{2j}} = 1$, $\sigma^2_{d_{3j}} = 1$, $\sigma^2_{c_j} = 2.245$, $\sigma^2_{a_j} = 1$, $\sigma^2_{c'_j} = 1$, and $\sigma^2_{b_j} = 1$. The mean of intercepts and covariances between random coefficients were set to zero (i.e., $d_1, d_2\ d_3$, σ_{d_{2j},a_j}, $\sigma_{d_{3j},c_j^{'}}$, σ_{d_{3j},b_j}, and $\sigma_{c_j^{'},b_j}$). As mentioned earlier, I used Equations 24 and 31 to calculate the population values for the inferred population parameters d_1, c, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, and σ^2_1. For example, $c = 0.35 \times 0.35 + 0.35 = 0.4725$

The simulation study used two design factors: the number of groups (clusters) and group (cluster) size. The number of groups was the number of Level-2 units while group size was the size of group at Level 1. All the groups had the same size; the multilevel data were balanced. The simulation was a 5 × 4 design. The number of groups took on 5 values: 50, 100, 200, 500, and 1000. The group size took on 4 values: 10, 20, 50, and 100. Within each condition, 1000 random data sets were generated based on multilevel mediation model in Equations 9 and 10. I used Equations 8-10 to estimate the parameters whose population values that were not set at zero. Note that Equation 8 is used to estimate inferred population parameters d_1, c, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, and σ^2_1. Note that I used balanced data in the simulation study to reduce effect of sampling errors on the sample estimates. It is expected that the effect of sampling errors on the sample estimates will be reduced as the number of clusters increases. The derived formulas use population values, not sample estimates with a sampling error. Thus to get close to the population values, a "large" number of clusters were needed to reduce the effect of sampling error on the estimates. Smaller numbers of clusters (e.g., 50) were also used in the simulation study because I did not know *a priori* what would constitute a "large" number of clusters.

Both data generation and estimation were performed in R. I wrote R code to generate data for the multilevel mediation model. For estimating the multilevel model, I used R package *lme4*. The outcome of the simulation study for each parameter were estimates of bias, relative bias, percentage bias, and *MSE*. The results are shown in Table 1 through Table 20.

To depict the results of the simulation more clearly, I created bubble charts to compare *MSE* for the inferred parameters d_1, c, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, and σ^2_1 (see Figure 3 through Figure 6). Hunt (2000) defined a bubble chart as "an extension of an *XY* scatterplot in which the size of the plotting symbol (a bubble) is used to display a third continuous variable" (p. 57). In each bubble chart, the *x*-axis is the group size and *y*-axis is the number of groups. The area of each bubble is proportional to the *MSE* of the respective parameter. In addition, in each figure, the *MSE* for the condition corresponding to the group size of 10 and the number of groups of 50 is shown. Note that the sizes of bubbles are the relative sizes of the *MSE*s used to compare the *MSE*s corresponding to various conditions for that specific parameter. The sizes of the bubbles are *not* comparable across figures.

Results. The simulation study indicates that the percentage bias for the inferred population values $c, \sigma^2_{c_j}, \sigma^2_1$ and $\sigma^2_{d_{1j}}$ got close to zero as the number of clusters increased. It appears that the derived formulas in Equations 24 and 31 for the inferred population parameters are correct.

Chapter 4

STUDY 2

Extending CWC2 to 1-1-1 Mediation Models

In this section, I extend the CWC2 centering strategy to capture between- and within-cluster effects. The centered model is termed the CWC2 1-1-1 mediation model. Using the CWC2 centering strategy, I demonstrate that the total effect of the predictor on the outcome variable (c), the effect of the predictor on the mediator (a), and the effect of the mediator on the outcome variable controlling for the predictor (b) can be decomposed into between- and within-cluster effects. The CWC2 strategy uses CWC scores as a Level-1 predictor and cluster means as a Level-2 predictor thereby separating the between- and within-person effects.

To obtain a CWC2 1-1-1 mediation model, one has to: (a) center each Level-1 predictor in Equations 4 - 7 using CWC, and (b) add cluster means as Level-2 predictors. The following are Level-1 equations for a CWC2 1-1-1 mediation model:

$$Y_{ij} = d_{1j} + c_j(X_{ij} - \bar{X}_j) + \varepsilon_{1ij} \tag{32}$$

$$M_{ij} = d_{2j} + a_j\ (X_{ij} - \bar{X}_j) + \varepsilon_{2ij} \tag{33}$$

$$Y_{ij} = d_{3j} + c_j^{'}(X_{ij} - \bar{X}_j) + b_j(M_{ij} - \bar{M}_j) + \varepsilon_{3ij} \tag{34}$$

The Level-2 (i.e., cluster-level) equations are as follows.

$$d_{1j} = d_1 + c_b\ \bar{X}_j + u_{d_{1j}}$$

$$d_{2j} = d_2 + a_b\ \bar{X}_j + u_{d_{2j}}$$

$$d_{3j} = d_3 + c_b^{'} \bar{X}_j + b_b \bar{M}_j + u_{d_{3j}}$$

$$c_j = c_w + u_{c_j}$$

$$a_j = a_w + u_{a_j}$$

$$c_j^{'} = c_w^{'} + u_{c_j^{'}}$$

$$b_j = b_w + u_{b_j}$$

where the subscript "*w*" denotes a *within-cluster* effect whereas the subscript "*b*" denotes a *between-cluster* coefficient.

The mixed model equations for the CWC2 1-1-1 mediation model are as follows:

$$Y_{ij} = d_1 + c_b \bar{X}_j + c_w (X_{ij} - \bar{X}_j) + u_{d_{1j}} + u_{c_j} (X_{ij} - \bar{X}_j) + \varepsilon_{1ij} \quad (35)$$

$$M_{ij} = d_2 + a_b \bar{X}_j + a_w (X_{ij} - \bar{X}_j) + u_{d_{2j}} + u_{a_j} (X_{ij} - \bar{X}_j) + \varepsilon_{2ij} \quad (36)$$

$$\begin{aligned} Y_{ij} = d_3 &+ c_b^{'} \bar{X}_j + b_b \bar{M}_j + c_w^{'} (X_{ij} - \bar{X}_j) + b_w (M_{ij} - \bar{M}_j) \\ &+ u_{d_{3j}} + u_{c^{'}_j} (X_{ij} - \bar{X}_j) + u_{b_j} (M_{ij} - \bar{M}_j) + \varepsilon_{3ij} \end{aligned} \quad (37)$$

where $\bar{X}_j$ and $\bar{M}_j$ are the observed means for cluster *j*. $(X_{ij} - \bar{X}_j)$ and $(M_{ij} - \bar{M}_j)$ are CWC scores for the independent and mediator variables, respectively. c_w is the total within-cluster effect, c_b is the total between-cluster effect, a_b is the between-cluster effect of X on *M*, a_w is the within-cluster effect of *X* on *M*, $c_b^{'}$ is the between-cluster direct effect, $c_w^{'}$ is the within-cluster direct effect, b_b is the between cluster effect of *M* on *Y* controlling for *X*, and b_w is the within-cluster effect of *M* on *Y* controlling for *X*. The parameters d_1, d_2, and d_3 are the intercepts. The terms $u_{d_{1j}}$, $u_{d_{2j}}$, $u_{d_{3j}}$, u_{c_j}, u_{a_j}, $u_{c'_j}$, and u_{b_j} are Level-2 *residuals* for the intercepts and slopes.

Statistical Assumptions in Correctly Specified CWC2 1-1-1 Mediation Models

Having specified the CWC2 multilevel mediation model in Equations 36 and 37, I discuss the statistical assumptions underlying a CWC2 1-1-1 mediation model. This section discusses the assumptions underlying a *correctly specified* CWC2 1-1-1 mediation model and the implications of the assumptions on the Level-2 covariance structure for that model.

I initially reformulate Equations 36 and 37 as matrices to make the assumptions more apparent:

$$\begin{aligned} &\text{Level-1:} \quad \mathbf{M}_j = \mathbf{X}_{2j}\, \boldsymbol{\beta}_{2j} + \boldsymbol{\varepsilon}_{2j} \\ &\text{Level-2:} \quad \boldsymbol{\beta}_{2j} = \mathbf{W}_{2j}\, \boldsymbol{\gamma}_2 + \mathbf{u}_{2j} \end{aligned} \tag{38}$$

$$\begin{aligned} &\text{Level-1:} \quad \mathbf{Y}_j = \mathbf{X}_{3j}\, \boldsymbol{\beta}_{3j} + \boldsymbol{\varepsilon}_{3j} \\ &\text{Level-2:} \quad \boldsymbol{\beta}_{3j} = \mathbf{W}_{3j}\, \boldsymbol{\gamma}_3 + \mathbf{u}_{3j} \end{aligned} \tag{39}$$

where,

$$\boldsymbol{\beta}_{2j} = \begin{pmatrix} d_{2j} \\ a_j \end{pmatrix}$$

$$\mathbf{X}_{2j} = \begin{pmatrix} \mathbf{1} & (\mathbf{X}_j - \mathbf{1}\,\bar{X}_j) \end{pmatrix}$$

$$\mathbf{W}_{2j} = \begin{pmatrix} 1 & \bar{X}_j & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{\gamma}_2 = \begin{pmatrix} d_2 \\ a_b \\ a_w \end{pmatrix}$$

$$\mathbf{u}_{2j} = \begin{pmatrix} u_{d_{2j}} \\ u_{a_j} \end{pmatrix}$$

$$\boldsymbol{\beta}_{3j} = \begin{pmatrix} d_{3j} \\ c_j^{'} \\ b_j \end{pmatrix}$$

$$\mathbf{X}_{3j} = \begin{pmatrix} \mathbf{1} & (\mathbf{X}_j - \mathbf{1}\bar{X}_j) & (\mathbf{M}_j - \mathbf{1}\bar{M}_j) \end{pmatrix}$$

$$\mathbf{W}_{3j} = \begin{pmatrix} 1 & \bar{X}_j & \bar{M}_j & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{\gamma}_3 = \begin{pmatrix} d_3 \\ c_b^{'} \\ b_b \\ c_w \\ b_w \end{pmatrix}$$

$$\mathbf{u}_{3j} = \begin{pmatrix} u_{d_{3j}} \\ u_{c_j^{'}} \\ u_{b_j} \end{pmatrix}$$

where $\mathbf{X}_j$, $\mathbf{M}_j$, and $\mathbf{Y}_j$ are $n_j \times 1$ vectors of the independent variable, mediator, and outcome variable for cluster *j*, respectively. $\mathbf{X}_{2j}$ is an $n_j \times 2$ matrix whose first column equals the constant one and the second column contains CWC scores, $(\mathbf{X}_j - \mathbf{1}\bar{X}_j)$. $\mathbf{X}_{3j}$ is an $n_j \times 3$ whose first column contains all ones, the second column contains CWC scores on the independent variable, $(\mathbf{X}_j - \mathbf{1}\bar{X}_j)$, and the third column contains the CWC scores on the mediator, $(\mathbf{M}_j - \mathbf{1}\bar{M}_j)$. $\mathbf{1}$ is an $n_j \times 1$ column vector of ones, and $\mathbf{1}\bar{X}_j$ and $\mathbf{1}\bar{M}_j$ are the column vectors of cluster means of the independent variable and mediator, respectively. $\boldsymbol{\beta}_{2j}$ and $\boldsymbol{\beta}_{3j}$ are Level-1 random coefficients. $\mathbf{W}_{2j}$ and $\mathbf{W}_{3j}$ are 2 × 3 and 3 × 5 matrices of Level-2 predictors, respectively. $\boldsymbol{\gamma}_2$ and $\boldsymbol{\gamma}_3$ are 3×1 and 5×1 Level-2 fixed effect coefficients, respectively. $\mathbf{u}_{2j}$ and $\mathbf{u}_{3j}$ are 3×1 and 5×1 column vectors of Level-2 residuals, respectively. $\boldsymbol{\varepsilon}_{2j}$ and $\boldsymbol{\varepsilon}_{3j}$ are $n_j \times 1$ column vectors of the Level-1 residuals.

The correct-specification assumption means that the CWC2 mediation model in Equations 36 and 37 represents the population's "true" causal model of the relationships between three variable *X*, *M*, and *Y* within each cluster. This assumption implies there is no omitted variable in the mediation model. In other words, given the CWC2 1-1-1 mediation model in Equations 36 and 37, the Level-1 and Level-2 residuals (errors of prediction) corresponding to the response variables $\mathbf{M}_j$ in Equations 36 and the ones associated with response variable $\mathbf{Y}_j$ in Equation 37 are independent from one another.

The implication this assumption for the Level-1 covariance structure is as follows. For either Equations 36 and 37, ε_{2ij} and ε_{3ij} are identically and independently distributed (i.i.d.) across individuals and clusters. That is, within each equation, conditional on the predictors as well as Level-2 residuals, Level-1 residuals are i.i.d. More importantly, the residual terms are independent across Equations 36 and 37. More succinctly stated, we have

$$\begin{pmatrix} \boldsymbol{\varepsilon}_{2j} \\ \boldsymbol{\varepsilon}_{3j} \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0}_{n_j \times 1} \\ \mathbf{0}_{n_j \times 1} \end{pmatrix}, \begin{pmatrix} \sigma_2^2 \, \mathbf{I}_{n_j} & \mathbf{0}_{n_j} \\ \mathbf{0}_{n_j} & \sigma_3^2 \, \mathbf{I}_{n_j} \end{pmatrix} \right) \tag{40}$$

where $\mathbf{0}_{n_j \times 1}$ is a zero column vector of size n_j, $\mathbf{0}_{n_j}$ is a zero square matrix of size n_j, and $\mathbf{I}_{n_j}$ is an identity square matrix of size n_j. Note that covariance between the Level-1 residuals across Equations 38 and 39 is zero.

In addition, let us examine the implication of correct specification assumption on the Level-2 covariance structure. For a correctly specified CWC2 1-1-1 model, Level-2 residuals in Equations 38 and 39 are assumed to be i.i.d with a multivariate normal distribution as follows:

$$\begin{pmatrix} \mathbf{u}_{2j} \\ \mathbf{u}_{3j} \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0}_{2\times 1} \\ \mathbf{0}_{3\times 1} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_{2,3} \\ \boldsymbol{\Sigma}_{2,3} & \boldsymbol{\Sigma}_3 \end{pmatrix} \right) \tag{41}$$

$$\boldsymbol{\Sigma}_2 = \begin{pmatrix} \sigma^2_{d_{2j}} & \sigma_{d_{2j},a_j} \\ \sigma_{d_{2j},a_j} & \sigma^2_{a_j} \end{pmatrix}$$

$$\boldsymbol{\Sigma}_3 = \begin{pmatrix} \sigma^2_{d_{3j}} & \sigma_{d_{3j},c'_j} & \sigma_{d_{3j},b_j} \\ \sigma_{d_{3j},c'_j} & \sigma^2_{c'_j} & \sigma_{c'_j,b_j} \\ \sigma_{d_{3j},b_j} & \sigma_{c'_j,b_j} & \sigma^2_{b_j} \end{pmatrix}$$

$$\boldsymbol{\Sigma}_{2,3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{42}$$

where $\mathbf{\Sigma}_2$ and $\mathbf{\Sigma}_3$ are the covariance matrices of $\mathbf{u}_{2j}$ ($\boldsymbol{\beta}_{2j}$) and $\mathbf{u}_{3j}$ ($\boldsymbol{\beta}_{3j}$), respectively. $\mathbf{\Sigma}_{2,3}$ is a 2×3 covariance[2] matrix between $\mathbf{u}_{2j}$ ($\boldsymbol{\beta}_{3j}$) and $\mathbf{u}_{3j}$ ($\boldsymbol{\beta}_{3j}$), where the rows correspond to the elements in the random vector $\mathbf{u}_{2j}$ ($\boldsymbol{\beta}_{3j}$) and the columns correspond to the elements in the residual vector $\mathbf{u}_{3j}$ ($\boldsymbol{\beta}_{3j}$).

Note that the correct specification assumption implies that the covariance matrix between Level-2 residuals $\mathbf{u}_{2j}$ and $\mathbf{u}_{3j}$ is zero (i.e., $\mathbf{\Sigma}_{2,3} = \mathbf{0}$). As a result, $\mathbf{u}_{2j}$ and $\mathbf{u}_{3j}$ are *independent* and have marginal multivariate normal distributions as follows (Anderson, 2003, Chapter 2):

$$\mathbf{u}_{2j} \sim N(\mathbf{0}_{2\times 1}, \mathbf{\Sigma}_2) \qquad (43)$$

$$\mathbf{u}_{3j} \sim N(\mathbf{0}_{3\times 1}, \mathbf{\Sigma}_3) \qquad (44)$$

Between-Cluster and Within-Cluster Mediated Effects

The between-cluster mediated effect is defined as the between-cluster effect of the predictor *X* on the response variable *Y* that is mediated by the mediator *M*. The product of two coefficients, $a_b\, b_b$, measures the between-cluster mediated effect. Note that the two coefficients are associated with $\bar{X}_j$ and $\bar{M}_j$, respectively. As mentioned earlier, $\bar{X}_j$ and $\bar{M}_j$ are the aggregated measures of X_{ij} and M_{ij} at the cluster level. Thus, the coefficients associated with $\bar{X}_j$ and

[2] Anderson (2003, p. 35) uses the term covariance for $\mathbf{\Sigma}_3$ as well as $\mathbf{\Sigma}_{2,3}$, whether the matrices are symmetric or not. Similarly, I will use the term covariance to refer to the covariance matrix of a vector as well as the covariance matrix between two vectors.

$\bar{M}_j$ measure the between-cluster effects involving the independent variable and mediator, respectively.

The within-cluster mediated effect is defined as the within-cluster effect of the independent variable *X* on the response variable *Y* that is mediated by the mediator *M*. The product of coefficients a_w and b_w quantifies the mediated effect. The two coefficients are associated with CWC scores of the independent variable and mediator, $(X_{ij} - \bar{X}_j)$ and $(M_{ij} - \bar{M}_j)$, respectively. The CWC scores measure the within-cluster effects associated with independent variable and mediator.

Example: Daily Diary Study

To clarify between- and within-cluster mediated effects, I will use the daily diary example presented earlier. In the daily diary study, the researcher is interested in studying the effect of daily stress on daily alcohol use. Applying CWC2, the researcher is able to estimate the within-person and between-person effect of stress on alcohol use. The within-person effect is the effect of daily stress on daily alcohol use while the between-person effect is the effect of chronic stress on chronic alcohol use. Next, the researcher is interested in conducting a multilevel mediation analysis. The researcher hypothesizes that negative mood (e.g., depression) mediates the effect of stress on alcohol use. In addition, the researcher is interested in testing whether there are between-person and within-person mediated effects. The two research questions that the researcher poses are as follows:

1. Within-person question: Does daily negative mood mediate the effect of daily stress on daily alcohol use?
2. Between-person question: Does chronic negative mood mediate the effect of chronic stress on alcohol use?

To answer the two questions, the researcher employs the CWC2 centering strategy. The CWC2 centering strategy, described earlier, allows the researcher to estimate the between- and within-person mediated effects (see Figure 2). For this example, the CWC2 1-1-1 mediation model is as follows:

$$Y_{ij} = d_1 + c_b \overline{STS}_j + c_w (STS_{ij} - \overline{STS}_j) + [\text{resid}] \tag{45}$$

$$M_{ij} = d_2 + a_b \overline{STS}_j + a_w (STS_{ij} - \overline{STS}_j) + [\text{resid}] \tag{46}$$

$$\begin{aligned} Y_{ij} = d_3 + c_b' \overline{STS}_j + b_b \overline{MOOD}_j + c_w' (STS_{ij} - \overline{STS}_j) \\ + b_w (MOOD_{ij} - \overline{MOOD}_j) + [\text{resid}] \end{aligned} \tag{47}$$

where $\overline{STS}_j$ is the person mean on stress, $(STS_{ij} - \overline{STS}_j)$ is the CWC score on stress, $\overline{MOOD}_j$ is the person mean on negative mood, and $(MOOD_{ij} - \overline{MOOD}_j)$ is the CWC score on negative mood. Note that $\overline{STS}_j$ and $\overline{MOOD}_j$ represent person-specific constructs that are different from the $(STS_{ij} - \overline{STS}_j)$ and $(MOOD_{ij} - \overline{MOOD}_j)$. $\overline{STS}_j$ and $\overline{MOOD}_j$ represent a person's *chronic stress* and *chronic negative mood* (chronic depression), respectively. $(STS_{ij} - \overline{STS}_j)$ and $(MOOD_{ij} - \overline{MOOD}_j)$ measure *daily fluctuations* in stress and negative mood scores. The person means and CWC scores contain *pure* between-person and within-person effects, respectively. In statistical terms, the person means and CWC scores are *orthogonal*. In other words, covariance between the person means and CWC scores are zero. As a result, the corresponding coefficient estimates have zero covariance.

The coefficient d_2 is the mean chronic alcohol use for a person whose chronic stress score is zero. If there is not an individual with chronic stress score equal to zero, then the chronic stress score should be centered. The coefficient a_b is the effect of chronic stress on chronic negative mood. The coefficient a_w is the effect of daily stress on daily negative mood. The coefficient d_3 is the mean chronic alcohol use for the individuals whose chronic stress and chronic negative mood scores are zero. c_b' is the direct effect of chronic stress on chronic alcohol use. b_b is the effect of chronic negative mood on chronic alcohol use controlling for chronic stress. c_w' is the direct effect of daily stress on daily alcohol use. b_w is the effect of daily negative mood on daily alcohol use controlling for daily stress.

The *between-person mediated effect* is the effect of chronic stress on chronic alcohol use that is transmitted through chronic negative mood. The effect is measured by the product of coefficients, $a_b b_b$. Chronic negative mood can be construed as a *between-person mediator*. In other words, the between-person mediator, chronic negative mood, mediates the between-person effect of an independent variable, chronic stress, on the response variable, chronic alcohol use. The within-person mediated effect is the effect of daily stress on daily alcohol use that is transmitted through daily negative mood. This effect is measured by $a_w \, b_w$. Daily negative mood is considered a *within-person mediator*. A within-person mediator (e.g., daily negative mood) mediates the within-person effect of independent variable (e.g., daily stress) on a response variable (e.g., daily alcohol use).

Relationship between Fixed and Random Effects in Correctly Specified CWC2 1-1-1 Mediation Models

The CWC2 centering strategy changes the expression $c = a\,b + c'$ as parameters in that formula are decomposed into between and within-cluster effects. Now, the question is what the formulas are for relationships between the between- (within-) cluster total effect, the between-(within-) cluster mediated effect and between-(within-) cluster direct effect in a CWC2 1-1-1 mediation model? This section presents analytical formulas that show the expression $c = a\,b + c'$ holds for between- and within-cluster fixed effects, separately. In addition, I derive relationships between the random effects in Equation 35 and the ones in Equations 36- 37 in the CWC2 1-1-1 mediation model. My strategy is to make Equations 35 and 37 comparable (i.e., contain the same covariates) by substituting the terms $(M_{ij} - \bar{M}_j)$ and $\bar{M}_j$ in (37) with the expressions that contains $(X_{ij} - \bar{X}_j)$ and $\bar{X}_j$. In other words, I reformulate Equation 37 so that it contains only $(X_{ij} - \bar{X}_j)$ and $\bar{X}_j$ as predictors.

I first obtain the formula for the mediator CWC score, $(M_{ij} - \bar{M}_j)$. To do so, it is convenient to derive the expression for the expected value of the mediator (i.e., cluster mean), $\bar{M}_j$. Given Level-2 (cluster specific) residual terms, the expected value of M_{ij} for each cluster is obtained by taking the expectation on all the observations within a cluster. Following is the formula for the cluster mean of the mediator:

$$\begin{aligned}\bar{M}_j &= \mathrm{E}[M_{ij} \mid \bar{X}_j, u_{d_{2j}}, u_{a_j}] \\ &= \mathrm{E}\left[d_2 + a_b\ \bar{X}_j + a_w\ (X_{ij} - \bar{X}_j) + u_{d_{2j}} + u_{a_j}(X_{ij} - \bar{X}_j) + \varepsilon_{2ij} \mid \bar{X}_j, u_{d_{2j}}, u_{a_j}\right] \\ &= d_2 + a_b\ \bar{X}_j + u_{d_{2j}}\end{aligned} \tag{48}$$

To obtain the expression for the mediator CWC score, $(M_{ij} - \bar{M}_j)$, I subtract the expression in (48) from the expression in (36). The result is follows:

$$\begin{aligned}M_{ij} - \bar{M}_j = &\left(d_2 + a_b\ \bar{X}_j + a_w\ (X_{ij} - \bar{X}_j) + u_{d_{2j}} + u_{a_j}(X_{ij} - \bar{X}_j) + \varepsilon_{2ij} \right) \\ &- d_2 + a_b\ \bar{X}_j + u_{d_{2j}} \\ &= a_w\ (X_{ij} - \bar{X}_j) + u_{a_j}(X_{ij} - \bar{X}_j) + \varepsilon_{2ij}\end{aligned} \tag{49}$$

Equation 49 is a multilevel model with two levels where the dependent variable is the CWC score for the mediator and the predictor is the CWC score for the independent variable. As can be seen, Equation 49 captures the unbiased estimate of the within-cluster relationship between X_{ij} and M_{ij}, denoted by a_w. In addition, Equation 49 takes into account both between- and within-cluster variation associated with u_{a_j} and ε_{2ij}, respectively. When I consider the analysis of multiple clusters, the $(M_{ij} - \bar{M}_j)$ s will be correlated within clusters (Neuhaus & Kalbfleisch, 1998). This within-cluster correlation is accounted for by u_{a_j} in (49).

Having obtained the expressions for $(M_{ij} - \bar{M}_j)$ and $\bar{M}_j$ in terms of $(X_{ij} - \bar{X}_j)$ and $\bar{X}_j$, I substitute the expression for $(M_{ij} - \bar{M}_j)$ in Equation 49 into Equation 37. To simplify the algebra, I first obtain the following three equations:

$$\begin{aligned} b_b \bar{M}_j &= b_b\left(d_2 + a_b \bar{X}_j + u_{d_{2j}}\right) \\ &= b_b\, d_2 + a_b\, b_b\, \bar{X}_j + b_b\, u_{d_{2j}} \end{aligned} \tag{50}$$

$$\begin{aligned} b_w\,(M_{ij} - \bar{M}_j) &= b_w\left(a_w (X_{ij} - \bar{X}_j) + u_{a_j}(X_{ij} - \bar{X}_j) + \varepsilon_{2ij}\right) \\ &= a_w\, b_w (X_{ij} - \bar{X}_j) + b_w\, u_{a_j}\,(X_{ij} - \bar{X}_j) + b_w\, \varepsilon_{2ij} \end{aligned} \tag{51}$$

$$\begin{aligned} u_{b_j}(M_{ij} - \bar{M}_j) &= u_{b_j}\left(a_w (X_{ij} - \bar{X}_j) + u_{a_j}(X_{ij} - \bar{X}_j) + \varepsilon_{2ij}\right) \\ &= a_w\, u_{b_j}(X_{ij} - \bar{X}_j) + u_{a_j}\, u_{b_j}(X_{ij} - \bar{X}_j) + u_{b_j}\, \varepsilon_{2ij} \end{aligned} \tag{52}$$

Substituting Equations 50- 52 into Equation 37, I arrive at the following expression:

$$\begin{aligned} Y_{ij} &= d_3 + c_b'\, \bar{X}_j + b_b\left(d_2 + a_b\, \bar{X}_j + u_{d_{2j}}\right) \\ &\quad + c_w'(X_{ij} - \bar{X}_j) + b_w\left(a_w (X_{ij} - \bar{X}_j) + u_{a_j}(X_{ij} - \bar{X}_j) + \varepsilon_{2ij}\right) \\ &\quad + u_{d_{3j}} + u_{c_j'}(X_{ij} - \bar{X}_j) + u_{b_j}\left(a_w (X_{ij} - \bar{X}_j) + u_{a_j}(X_{ij} - \bar{X}_j) + \varepsilon_{2ij}\right) + \varepsilon_{3ij} \\ &= d_3 + c_b'\, \bar{X}_j + b_b\, d_2 + a_b\, b_b\, \bar{X}_j + b_b\, u_{d_{2j}} + c_w'(X_{ij} - \bar{X}_j) \\ &\quad + a_w\, b_w\,(X_{ij} - \bar{X}_j) + b_w\, u_{a_j}(X_{ij} - \bar{X}_j) + b_w\, \varepsilon_{2ij} + u_{d_{3j}} \\ &\quad + u_{c_j'}(X_{ij} - \bar{X}_j) + a_w\, u_{b_j}(X_{ij} - \bar{X}_j) + u_{a_j} u_{b_j}(X_{ij} - \bar{X}_j) + u_{b_j}\, \varepsilon_{2ij} + \varepsilon_{3ij} \\ &= (d_3 + b_b\, d_2) + (c_b' + a_b\, b_b)\bar{X}_j + (c_w' + a_w\, b_w)(X_{ij} - \bar{X}_j) \\ &\quad + (b_b\, u_{d_{2j}} + u_{d_{3j}}) + (u_{c_j'} + b_w\, u_{a_j} + a_w\, u_{b_j} + u_{a_j}\, u_{b_j})\,(X_{ij} - \bar{X}_j) \\ &\quad + \left((u_{b_j} + b_w)\,\varepsilon_{2ij} + \varepsilon_{3ij}\right) \end{aligned}$$

Thus, the following holds:

$$
\begin{aligned}
&Y_{ij} = (d_3 + b_b\, d_2) + (c_b^{'} + a_b\, b_b)\, \bar{X}_j + (c_w^{'} + a_w\, b_w)\,(X_{ij} - \bar{X}_j) + \\
&(u_{d_{3j}} + b_b\, u_{d_{2j}}) + \left(u_{c_j^{'}} + a_w\, u_{b_j} + (b_w + u_{b_j})\, u_{a_j}\right)(X_{ij} - \bar{X}_j) + \left((u_{b_j} + b_w)\, \varepsilon_{2ij} + \varepsilon_{3ij}\right)
\end{aligned}
\tag{53}
$$

Equation 53 reformulates Equation 37 so that it contains only $(X_{ij} - \bar{X}_j)$ and $\bar{X}_j$ as predictors. Now Equation 35 and Equation 53 are comparable.

To obtain algebraic relationships between the fixed effects in Equations 35 and the ones in Equation 53, I first obtain the expected values of both equations as follows:

$$
\mathrm{E}(Y_{ij} \mid X_{ij}) = d_1 + c_b\, \bar{X}_j + c_w\,(X_{ij} - \bar{X}_j) \tag{54}
$$

$$
\begin{aligned}
\mathrm{E}(Y_{ij} \mid X_{ij}) = {} & (d_3 + b_b\, d_2) + (c_b^{'} + a_b\, b_b)\, \bar{X}_j \\
& + (c_w^{'} + a_w\, b_w)\,(X_{ij} - \bar{X}_j)
\end{aligned}
\tag{55}
$$

Because Equations 54 and 55 are equal on the left side, the corresponding terms on the right side (i.e., intercepts and slopes) must be equal for all $\bar{X}_j$ s and $(X_{ij} - \bar{X}_j)$ s , or all intercepts and the slopes must be equal to zero. Because the intercepts and slopes are not zero, the corresponding intercepts and slopes from each equation are equal. That is, the following equalities hold:

$$
\begin{aligned}
d_1 &= d_3 + b_b\, d_2 \\
c_b &= c_b^{'} + a_b\, b_b \\
c_w &= c_w^{'} + a_w\, b_w
\end{aligned}
\tag{56}
$$

Equation 56 shows that the CWC2 centering strategy decomposes the total effect of the predictor on the outcome variable (c), the effect of the predictor on the mediator (a), and the effect of the mediator on the outcome variable controlling for the predictor (b) into the between- and within-cluster effects. The result in (56) indicates that the between-cluster total effect (i.e., c_b) is equal to the between-cluster mediated effect (i.e., $a_b\ b_b$) plus the between-cluster direct effect (i.e., $c_b^{'}$). Similarly, the within-cluster total effect (i.e., c_w) is equal to the within-cluster mediated effect (i.e., $a_w\ b_w$) plus the within-cluster direct effect (i.e, $c^{'}_w$).

Further, I can obtain the relationships between Level-1 and Level-2 residual terms. To do so, first I obtain the residuals for Equations 35 and 53 , denoted by r_1 and r_1^*, respectively, using $Y_{ij} - \mathrm{E}(Y_{ij} \mid X_{ij})$. The residuals r_1 and r_1^* are shown below:

$$r_1 = u_{d_{1j}} + u_{c_j}(X_{ij} - \bar{X}_j) + \varepsilon_{1ij} \tag{57}$$

$$\begin{aligned} r_1^* &= (u_{d_{3j}} + b_b\, u_{d_{2j}}) + \left(u_{c_j^{'}} + a_w\, u_{b_j} + (b_w + u_{b_j})\, u_{a_j}\right)(X_{ij} - \bar{X}_j) \\ &\quad + \left((u_{b_j} + b_w)\, \varepsilon_{2ij} + \varepsilon_{3ij}\right) \end{aligned} \tag{58}$$

Note that r_1 and r_1^* contain both Level-1 and Level-2 residual terms. To isolate Level-2 residuals, I utilize the fact that Level-2 residual terms are fixed at Level 1 (i.e., within cluster). By taking within-cluster expectation, I can obtain Level-2 residual terms as follows:

$$\begin{aligned} u_1 &= \mathrm{E}(r_1 \mid u_{c_j}, u_{d_{1j}}, X_{ij}) \\ &= \mathrm{E}(u_{d_{1j}} + u_{c_j}(X_{ij} - \bar{X}_j) + \varepsilon_{1ij} \mid u_{c_j}, u_{d_{1j}}, X_{ij}) \\ &= u_{d_{1j}} + u_{c_j}\,(X_{ij} - \bar{X}_j) \end{aligned} \tag{59}$$

$$
\begin{aligned}
u_1^* &= \mathrm{E}(r_1^* \mid u_{d_{2j}}, u_{a_j}, u_{d_{3j}}, u_{c_j'}, u_{b_j}, X_{ij}) \\
&= \mathrm{E}\Big[(u_{d_{3j}} + b_b\, u_{d_{2j}}) + \big(u_{c_j'} + a_w\, u_{b_j} + (b_w + u_{b_j})\, u_{a_j}\big)(X_{ij} - \bar{X}_j) \\
&\quad + \big((u_{b_j} + b_w)\, \varepsilon_{2ij} + \varepsilon_{3ij}\big) \mid u_{d_{2j}}, u_{a_j}, u_{d_{3j}}, u_{c_j'}, u_{b_j}, X_{ij}\Big] \\
&= (u_{d_{3j}} + b_b\, u_{d_{2j}}) + \big(u_{c_j'} + a_w\, u_{b_j} + (b_w + u_{b_j})\, u_{a_j}\big)(X_{ij} - \bar{X}_j) \\
&\quad + u_{b_j}\, \mathrm{E}(\varepsilon_{2ij}) + b_w\, \mathrm{E}(\varepsilon_{2ij}) + \mathrm{E}(\varepsilon_{3ij}) \\
&= (u_{d_{3j}} + b_b\, u_{d_{2j}}) + \big(u_{c_j'} + a_w\, u_{b_j} + (b_w + u_{b_j})\, u_{a_j}\big)(X_{ij} - \bar{X}_j)
\end{aligned} \tag{60}
$$

where u_1 and u_1^* are Level-2 residuals. As can be seen, the term associated with $(X_{ij} - \bar{X}_j)$ is the slope residual while the remainder is the intercept residual. Because Equations 59 and 60 are equal on the left side of equal sign, the corresponding residual terms on the right side must also be equal for all $(X_{ij} - \bar{X}_j)$, or they all must be equal to zero. Because the residual terms are not zero the corresponding terms in each equation are equal. That is,

$$
\begin{aligned}
u_{d_{1j}} &= u_{d_{3j}} + b_b\, u_{d_{2j}} \\
u_{c_j} &= u_{c_j'} + b_w\, u_{a_j} + a_w\, u_{b_j} + u_{a_j}\, u_{b_j}
\end{aligned} \tag{61}
$$

Finally, to obtain the relationship between Level-1 residuals, I first derive the expressions for the Level-1 residuals in Equations 35 and 53 by subtracting u_1 and u_1^* from r_1 and r_1^*, respectively. Let e_1 and e_1^* denote Level-1 residuals. Then,

$$e_1 = r_1 - u_1 = \varepsilon_{1ij}$$

$$e_1^* = r_1^* - u_1^* = (u_{b_j} + b_w)\, \varepsilon_{2ij} + \varepsilon_{3ij}$$

Thus, the following holds:

$$\varepsilon_{1ij} = (u_{b_j} + b_w)\, \varepsilon_{2ij} + \varepsilon_{3ij} \tag{62}$$

Equations 61 and 62 show the relationships between Level-1 and Level-2 residual terms in Equation 35 and the Level-1 and Level-2 residual terms in Equations 36-37.

Further, I can use the expressions in (61) and (62) to obtain the variances for d_{1j}, c_j and ε_{1ij} and the covariance between d_{1j} and c_j in terms of the variances and covariances of the residual terms in Equations 36 and 37. Below is the detailed description of the analytic derivations for the variances of d_{1j}, c_j and ε_{1ij} and the covariance between i_{1j} and c_j. First, I derive the variance for d_{1j} as follows:

$$\begin{aligned}
\sigma^2_{d_{1j}} &= \mathrm{var}(u_{d_{3j}} + b_b\, u_{d_{2j}}) \\
&= \mathrm{var}(u_{d_{3j}}) + b_b^{\,2}\ \mathrm{var}(u_{d_{2j}}) + 2\ b_b\, \mathrm{cov}(u_{d_{3j}}, u_{d_{2j}}) \\
&= \sigma^2_{d_{3j}} + b_b^{\,2}\sigma^2_{d_{2j}}
\end{aligned}$$

Note that $Cov(u_{b_j}, u_{d_{2j}}) = 0$ because u_{b_j} and $u_{d_{2j}}$ are independent.

The variance of c_j is obtained as follows:

$$\begin{aligned}
\sigma^2_{c_j} = {} & \mathrm{var}(u_{c_j'}) + \mathrm{var}(b_w\, u_{a_j}) + \mathrm{var}(a_w\, u_{b_j}) + \mathrm{var}(u_{a_j}\, u_{b_j}) \\
& + 2\left\{ a_w\ \mathrm{cov}(u_{c_j'}, u_{b_j}) + b_w\, \mathrm{cov}(u_{c_j'}, u_{a_j}) + \mathrm{cov}(u_{c_j'}, u_{a_j} u_{b_j}) \right. \\
& \left. + \mathrm{cov}(a_w\, u_{b_j}, b_w\, u_{a_j}) + \mathrm{cov}(a_w\, u_{b_j}, u_{a_j} u_{b_j}) + \mathrm{cov}(b_w\, u_{a_j}, u_{a_j} u_{b_j}) \right\} \\
= {} & \sigma^2_{c_j'} + a_w^{\,2}\sigma^2_{b_j} + (b_w^{\,2} + \sigma^2_{b_j})\sigma^2_{a_j} + 2\, a_w\, \sigma_{c_j', b_j}
\end{aligned}$$

Note that in deriving the above equation, the following relations hold:

$$\begin{aligned}
\mathrm{cov}(u_{c_j'}, u_{a_j}\, u_{b_j}) &= \mathrm{E}(u_{c_j'}\, u_{a_j}\, u_{b_j}) \\
&= \mathrm{E}(u_{a_j}\, u_{c_j'}\, u_{b_j}) \\
&= \mathrm{E}(u_{a_j})\, \mathrm{E}(u_{c_j'}\, u_{b_j}) \\
&= 0
\end{aligned}$$

$$\begin{aligned}\operatorname{cov}(a_w\, u_{b_j}, b_w\, u_{a_j}) &= a_w\, b_w\ \operatorname{cov}(u_{b_j}, u_{a_j})\\ &= 0\end{aligned}$$

$$\begin{aligned}\operatorname{cov}(a_w\, u_{b_j}, u_{a_j} u_{b_j}) &= a_w\ \operatorname{cov}(u_{b_j}, u_{a_j}\, u_{b_j})\\ &= a_w\, \mathrm{E}(u_{a_j}\, u_{b_j}^2)\\ &= a_w\, \mathrm{E}(u_{a_j})\, \mathrm{E}(u_{b_j}^2)\\ &= 0\end{aligned}$$

$$\begin{aligned}\operatorname{cov}(b_w\, u_{a_j}, u_{a_j}\, u_{b_j}) &= b_w \operatorname{cov}(u_{a_j}, u_{a_j}\, u_{b_j})\\ &= b_w\, \mathrm{E}(u_{a_j}^2\, u_{b_j})\\ &= b_w\, \mathrm{E}(u_{a_j}^2)\, \mathrm{E}(u_{b_j})\\ &= 0\end{aligned}$$

In addition, for Level-1 variance, σ_1^2, we have

$$\begin{aligned}\sigma_1^2 &= \operatorname{var}(\varepsilon_{1ij}) = \operatorname{var}(b_w\, \varepsilon_{2ij} + u_{b_j}\, \varepsilon_{2ij} + \varepsilon_{3ij})\\ &= \operatorname{var}(b_w\, \varepsilon_{2ij}) + \operatorname{var}(u_{b_j}\, \varepsilon_{2ij}) + \operatorname{var}(\varepsilon_{3ij})\\ &\quad +2\left\{\operatorname{cov}(b_w\, \varepsilon_{2ij}, u_{b_j}\, \varepsilon_{2ij}) + \operatorname{cov}(b_w\, \varepsilon_{2ij}, \varepsilon_{3ij}) + \operatorname{cov}(u_{b_j}\, \varepsilon_{2ij} + \varepsilon_{3ij})\right\}\\ &= b_w^2\, \mathrm{E}(\varepsilon_{2ij}^2) + \mathrm{E}(u_{b_j}^2 \varepsilon_{2ij}^2) + \mathrm{E}(\varepsilon_{3ij}^2)\\ &\quad +2\left\{b_w\, \mathrm{E}(\varepsilon_{2ij}\, u_{b_j}\, \varepsilon_{2ij}) + b_w\, \mathrm{E}(\varepsilon_{2ij}\, \varepsilon_{3ij}) + \mathrm{E}(u_{b_j}\, \varepsilon_{2ij}\, \varepsilon_{3ij})\right\}\\ &= b_w^2\, \sigma_2^2 + \sigma_{b_j}^2\, \sigma_2^2 + \sigma_3^2 + 2\,\{0+0+0\}\\ &= (b_w^2 + \sigma_{b_j}^2)\, \sigma_2^2 + \sigma_3^2\end{aligned}$$

I can also use the two equations in (61) to derive the covariance between u_{c_j} and $u_{d_{1j}}$ as follows:

$$
\begin{aligned}
\sigma_{u_{d_{1j}},u_{c_j}} &= \mathrm{E}\left[(u_{d_{1j}} + b_b\, u_{d_{2j}})(u_{c_j^{'}} + b_w\, u_{a_j} + a_w\, u_{b_j} + u_{a_j}\, u_{b_j})\right] \\
&= \mathrm{E}\big[u_{d_{3j}}\, u_{c_j^{'}} + b_w\, u_{d_{3j}}\, u_{a_j} + a_w\, u_{d_{3j}}\, u_{b_j} + u_{d_{3j}}\, u_{a_j}\, u_{b_j} \\
&\quad + b_b\, u_{d_{2j}}\, u_{c_j^{'}} + b_b\, u_{d_{2j}}\, b_w\, u_{a_j} + a_w\, b_b\, u_{d_{2j}}\, u_{b_j} + b_b\, u_{d_{2j}}\, u_{a_j} u_{b_j}\big] \\
&= \mathrm{E}(u_{d_{3j}}\, u_{c_j^{'}}) + \mathrm{E}(b_w\, u_{d_{3j}}\, u_{a_j}) + \mathrm{E}(a_w\, u_{d_{3j}}\, u_{b_j}) + \mathrm{E}(u_{d_{3j}}\, u_{a_j} u_{b_j}) \\
&\quad + \mathrm{E}(b_b\, u_{d_{2j}}\, u_{c_j^{'}}) + \mathrm{E}(b_b\, b_w\, u_{d_{2j}}\, u_{a_j}) + \mathrm{E}(a_w\, b_b\, u_{d_{2j}}\, u_{b_j}) + \mathrm{E}(b_b\, u_{d_{2j}}\, u_{a_j} u_{b_j}) \\
&= \mathrm{E}(u_{d_{3j}}\, u_{c_j^{'}}) + b_w\, \mathrm{E}(u_{d_{3j}}\, u_{a_j}) + a_w\, \mathrm{E}(u_{d_{3j}}\, u_{b_j}) + \mathrm{E}(u_{d_{3j}}\, u_{a_j} u_{b_j}) \\
&\quad + b_b\, \mathrm{E}(u_{d_{2j}}\, u_{c_j^{'}}) + b_b\, b_w\, \mathrm{E}(u_{d_{2j}}\, u_{a_j}) + a_w\, b_b\, \mathrm{E}(u_{d_{2j}}\, u_{b_j}) + b_b\, \mathrm{E}(u_{d_{2j}}\, u_{a_j}\, u_{b_j}) \\
&= \sigma_{u_{d_{2j}},c_j^{'}} + 0 + a_w\, \sigma_{u_{d_{3j}},u_{b_j}} + 0 + 0 + b_b\, b_w\, \sigma_{u_{d_{2j}},u_{a_j}} + 0 + 0 \\
&= \sigma_{u_{d_{3j}},c_j^{'}} + a_w\, \sigma_{u_{d_{3j}},u_{b_j}} + b_b\, b_w\, \sigma_{u_{d_{2j}},u_{a_j}} \\
&= \sigma_{d_{3j},c_j^{'}} + a_w\, \sigma_{d_{3j},b_j} + b_b\, b_w\, \sigma_{d_{2j},a_j}
\end{aligned}
$$

Below is the summary of the derived relationships between the variances and covariances of Level-1 and Level-2 residuals:

$$
\begin{aligned}
\sigma^2_{d_{1j}} &= \sigma^2_{d_{3j}} + b_b^2\, \sigma^2_{d_{2j}} \\
\sigma^2_{c_j} &= \sigma^2_{c_j^{'}} + a_w^2\, \sigma^2_{b_j} + (b_w^2 + \sigma^2_{b_j})\, \sigma^2_{a_j} + 2\, a_w\, \sigma_{c_j^{'},b_j} \\
\sigma^2_1 &= (b_w^2 + \sigma^2_{b_j})\, \sigma^2_2 + \sigma^2_3 \\
\sigma_{d_{1j},c_j} &= \sigma_{d_{3j},c_j^{'}} + a_w\, \sigma_{d_{3j},b_j} + b_b\, b_w\, \sigma_{d_{2j},a_j}
\end{aligned}
\qquad (63)
$$

Equations 56 and 63 show the relationships between the fixed effects and the variances and covariances of random effects in Equation 35 and the fixed effects and the variances and covariances of random effects in Equations 36 and 37.

Simulation Study

This section describes a small-scale simulation study used to assess whether the derived relationships in Equations 56 and 63 are empirically accurate. In this study, I used the derived relationships in Equations 56 and 63 to calculate the population values for the parameters d_1, c_w, c_b, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, σ^2_1, and σ_{d_{1j},c_j}. Note that these parameters are not free parameters. They are a function of other parameters as shown in Equations 36 and 37. These parameters are referred to as *inferred* parameters as indicated in Kenny et al. (2003).

The inferred population values are then used to estimate bias, relative bias, and mean squared error (*MSE*). For example, for the inferred parameter d_1, the estimate ($\hat{d}_1$) , the bias, relative bias, and *MSE* are obtained as follows:

$$\hat{d}_1 = \frac{\sum \hat{d}_1^{rep}}{n}$$

$$bias = \hat{d}_1 - d_1$$

$$relative\,bias = \frac{\hat{d}_1 - d_1}{d_1}$$

$$MSE = \frac{\sum (\hat{d}_1^{rep} - d_1)^2}{n-1}$$

where $\hat{d}_1^{rep}$ is the REML estimate for d_1 obtained from each simulation replication *rep,* and *n* is total the number of replications. Inferred parameter d_1 is obtained from Equation 56.

Simulation Design. The population model for the simulation study was based on the CWC2 1-1-1 mediation model in Equations 36 and 37. The population parameters for the simulation are as follows: $d_1 = 0.575$, $d_2 = 0.5$, $d_3 = 0.5$, $c_w = 0.4724$, $c_b = 0.1725$, $a_w = 0.35$, $a_b = 0.15$, $c_w^{'} = 0.35$, $c_b^{'} = 0.15$, $b_w = 0.35$, $b_b = 0.15$, $\sigma_1^2 = 2.1225$, $\sigma_2^2 = 1$, $\sigma_3^2 = 1$, $\sigma^2_{d_{1j}} = 1.0225$, $\sigma^2_{d_{2j}} = 1$,

$\sigma^2_{d_{3j}} = 1$, $\sigma^2_{c_j} = 2.315$, $\sigma^2_{a_j} = 1$, $\sigma^2_{c'_j} = 1$, $\sigma^2_{b_j} = 1$, $\sigma_{c'_j, b_j} = 0.1$, $\sigma_{d_{2j}, a_j} = 0.1$, $\sigma_{d_{3j}, c'_j} = 0.1$, $\sigma_{d_{3j}, b_j} = 0.1$ and σ_{d_{1j}, c_j} =0.14025. As mentioned earlier, I used Equations 56 and 63 to calculate the population values for the inferred parameters d_1, c_w, c_b, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, σ^2_1, and σ_{d_{1j}, c_j}. For example, c_w = 0.35 × 0.35 + 0.35 = 0.4725.

The simulation study used one design factor: the number of groups (clusters) and group (cluster) size. The number of groups was the number of Level-2 units while group size was the size of the groups at Level 1. All of the groups have the same size; the multilevel data are balanced. The simulation was a 5 × 1 design. The number of groups took on 5 values: 50, 100, 200, 500, and 1000. The group size took on one value: 10. I chose the group size of 10 because the Study 1 simulation results showed that Level-1 sample size did not have a substantial effect on the outcome of simulation. Within each condition, 1000 random data sets were generated based on the CWC2 multilevel mediation model in Equations 36 and 37. I used Equations 56 and 63 to estimate the parameters of interest. As mentioned above, I used balanced data in the simulation study to reduce effect of sampling errors on the sample estimates. I expected that the effect of sampling error on the sample estimates would be reduced as the number of clusters increased. The derived formulas use population values, not sample estimates with a sampling error. Thus to get close to the population values, a "large" number of clusters were needed to reduce the effect of sampling error on the estimates. Smaller number of clusters (e.g., 50) was also used in the simulation study because I did not know *a priori* what would constitute a "large" number of clusters.

Both data generation and estimation were performed in R. I wrote R code to generate data for the CWC2 1-1-1 mediation model. For estimating the multilevel model, I used R package *lme4*. The outcomes of the simulation study for each parameter were estimates of bias, relative bias, percentage bias, and *MSE*. The results are shown in Table 21 - 25.

Results. The simulation study indicates that the percentage bias for the derived population values for the inferred parameters d_1, c_w, c_b, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, σ^2_1, and σ_{d_{1j},c_j} got close to zero as the number of clusters increased. It appears that the derived formulas in Equations 56 and 63 for the inferred population parameters to be correct.

Chapter 5

CONCLUSIONS AND DISCUSSION

Many prevention interventions occur in cluster settings (classroom or community) or involve repeated measures (multiple measurements of drug use) over time giving rise to a need for mediation models that take clustering into account. Research on multilevel mediation analysis is relatively recent, and several issues warrant more attention. In this dissertation, I focused on two issues. The first issue was the ambiguity in specifying the underlying assumptions in 1-1-1 mediation models. It appears that there is no current consensus among researchers on the assumptions needed for 1-1-1 mediation models. These assumptions have different consequential implications for the specification of the Level-2 covariance structure. The second issue is the effect of the centering strategy on algebraic relationships between the quantities of interest such as the total effect, mediated effect, and direct effect. To the best of my knowledge, given CWC2 centering, no research to date has presented formulas for the relationships between between-cluster (within-cluster) total effect, between-cluster (within-cluster) mediated effect, and between-cluster (within-cluster) direct effect.

Study 1

Statistical Assumptions in Correctly Specified Uncentered 1-1-1 Mediation Models. I discussed the assumptions underlying a correctly specified uncentered 1-1-1 mediation model and the implications of those assumptions on the specification of the Level-2 covariance structure. If an uncentered 1-1-1 mediation model is correctly specified, this implies that there are no omitted variables. This result implies that residuals cannot be correlated. In other words, in a correctly specified uncentered 1-1-1 mediation model, Level-1 and Level-2 residuals may *not* be correlated across Equations 5 and 6.

Examining Kenny et al.'s (2003) **and Bauer et al.'s** (2006) **Approach**. Kenny et al. (2003) and Bauer et al. (2006) assumed that the random coefficients a_j and b_j are correlated across Equations 5 and 6. Note that the correlation between a_j and b_j implies that the associated Level-2 residuals (i.e., $u_{d_{2j}}$ and $u_{d_{3j}}$) are also correlated. As mentioned previously, correlated residuals across equations implies that the model is *misspecified*. In other words, Kenny et al.'s (2003) postulated uncentered 1-1-1 mediaion model is *not* correctly specified.

Relationships between parameters in correctly specified uncentered 1-1-1 mediation models. For a correctly specified 1-1-1 mediation model, I presented algebraic relationships between the fixed and random effects in Equation 8 and the fixed and random effects in Equations 9 and 10. Note that these results are only valid if the underlying assumptions are met. Of the most importance for the relationships that were derived is the expression $c = a\,b + c'$. That is, the total effect is equal to the mediated effect plus the direct effect. To the best of my knowledge, no previous study has provided an analytical proof for this algebraic relationship for the multilevel case. Note that all of the parameters in this expression are "true" population values, not sample estimates. Krull and MacKinnon (2001) examined this expression using a simulation study but they did not provide a mathematical proof of this expression. Krull and MacKinnon concluded that this expression may not hold when using ML estimates. That is, $\hat{c} \neq \hat{a}\hat{b} + \hat{c}'$. The discrepancy in Krull and MacKinnon's simulation result could be a result of the sampling error of the estimates, especially for the estimator of the parameter *c*. To verify my derived relationships hold for the population values, I conducted a simulation study when the number of clusters ranged from 50 to 1000. When the cluster size is large (e.g., 500 or 1,000), these relationships hold because the sample estimates became closer to the population values and the sampling error became smaller. In other words, to verify the expression in a simulation study, the number of clusters is required to be "large" enough to mitigate the effect of sampling errors on the ML estimates of the parameters. The result of the simulation study revealed that expression $c = ab + c'$ holds when the number of clusters becomes "large" (e.g., 1000).

Study 2

Extending the CWC2 centering strategy to CWC2 1-1-1 mediation models. I extended the CWC2 centering strategy to a correctly specified (i.e., no correlated residuals) 1-1-1 mediation model. CWC2 is a centering strategy that uses CWC scores as a Level-1 predictor and cluster means as a Level-2 predictor thereby separating the between- and within-person effects. In 1-1-1 mediation models, CWC2 is implemented as follows: a) center each Level-1 predictor in Equations 4 - 7 using CWC, and b) add the cluster mean as a Level-2 predictor of the random intercept and slope. I use the term CWC2 1-1-1 mediation to describe a 1-1-1 mediation model that used the CWC2 centering strategy. Equations 35- 37 specify a CWC2 1-1-1 mediation model.

Statistical assumptions in CWC2 1-1-1 mediation models. I discussed some of the assumptions underlying a correctly specified CWC2 1-1-1 mediation model. A correctly specified CWC2 1-1-1 mediation model is one without correlated residuals across regression equations. More specifically, Level-1 residuals and Level-2 residuals may *not* be correlated across Equations 36 and 37.

The CWC2 centering strategy was applied to the correctly specified, uncentered 1-1-1 mediation model. However, the CWC2 centering strategy can also be applied to the misspecified uncentered 1-1-1 mediation model expressed by Equations 8 - 10. The equations for the misspecified CWC2 multilevel mediation model were the same as the equations for the correctly specified model in 35- 37 except for the nonzero covariance between the Level-2 residuals u_{a_j} and u_{b_j}. This result also implies that a_j and b_j will be correlated and thus $\sigma_{u_{aj},u_{bj}} = \sigma_{a_j,b_j}$. An implication of this argument is that the CWC2 centering strategy makes the between- and within-cluster effects orthogonal; however, CWC2 does *not* account for the Level 2 correlation between Level-2 residuals. To illustrate, I analyzed the data from the example presented by Kenny et al. (2003) using Mplus (L. K. Muthén & B. O. Muthén, 1998-2010). The analysis of the misspecified uncentered 1-1-1 mediation model with correlated residuals yielded an estimate of the covariance $\sigma_{\hat{a},\hat{b}} = 0.126$. The analysis of the misspecified CWC2 1-1-1 mediation model in which a_j and b_j were correlated yielded an estimate of covariance equal to 0.131. This example indicates that the

CWC2 centering strategy *cannot* account for the correlation between Level-2 residuals.

Between-cluster and within-cluster mediated effects. I presented definitions for the between-cluster and within-cluster mediated effects as well as the mediators. A between-cluster mediator is a Level-2 intervening variable that mediates the between-cluster effect of a predictor on an outcome. Note that in a CWC2 1-1-1 mediation model, the between-cluster mediator is the cluster mean of the mediator and the predictor is cluster mean of the independent variable. A within-cluster mediator is an intervening variable that contains only within-cluster variation and mediates the within-cluster effect of a predictor on the within-cluster effect of an outcome. In a CWC2 1-1-1 mediation model, a within-cluster mediator is the CWC score on the mediator; the predictor is the CWC score on the independent variable. Using the example of daily diary study of stress-negative mood-alcohol use, the between-person mediated effect is the effect of chronic stress on chronic alcohol use transmitted through chronic negative mood. The between-person (-cluster) mediated effect is measured by $a_b\ b_b$. Chronic negative mood is the between-person mediator. Similarly, the within-person mediated effect is the effect of the deviation from the mean level of daily stress on the deviation from the mean level of daily alcohol use mediated by deviation from the mean level of daily negative mood. The within-person mediated effect is quantified by $a_w\ b_w$. The deviation of the mean level of daily negative mood, measured by the CWC score on daily negative, is a within-person mediator.

Relationship between fixed and random effects in CWC2 1-1-1 mediation models. For a correctly specified CWC2 1-1-1 mediation model, I presented algebraic relationships between the fixed and random effects in Equation 35 and the fixed and random effects in Equations 36 and 37. Note that these results depict the relationships between the population parameters and are valid only if the underlying assumptions are met. Of the most importance are the expressions $c_b = a_b\, b_b + c_b^{'}$ and $c_w = a_w\, b_w + c_w^{'}$. The expressions indicate that the between-cluster total effect (i.e., c_b) is equal to the between-cluster mediated effect (i.e., $a_b\, b_b$) plus the between-cluster direct effect (i.e., $c_b^{'}$). Similarly, the within-cluster total effect (i.e., c_w) is equal to the within-cluster mediated effect (i.e., $a_w\, b_w$) plus the within-cluster direct effect (i.e, $c^{'}_{w}$). To the best of my knowledge, no previous study has derived these algebraic relationships.

Limitations

One limitation of the current study is that in Equation 48 the expected value of the mediator instead of the observed cluster mean is used to derive algebraic relationships in a CWC2 multilevel mediation model. While this may not affect the results concerning the relationships between the fixed effects, it can affect the derived relationships between the random effects. In practice, however, researchers are more interested in the relationships between the fixed effects. The use of the expected value (true cluster mean) instead of observed cluster mean does not change the derived relationships between the fixed effects.

Another limitation of this study is that two-level 1-1-1 mediation models were discussed. The multilevel mediation models that were not considered in the present study are as follows: 2-2-1, 2-1-1, 2-1-2, 1-2-1, 1-1-2, and 1-2-2. Note that in these two-level mediation models, the within-cluster mediated effect cannot be estimated because at least one of the variables in the model (*X, M,* or *Y*) will *not* vary within clusters. In these models, only between-cluster mediated effects can be estimated (Zhang et al., 2009). In addition, the results of the present study are limited to the multilevel mediation models in which variables are measured at two levels. Multilevel mediation models with three levels can also be considered. However, it is not clear how the CWC2 centering strategy can be extended to the multilevel mediation models with more than two levels.

In addition, in the current study, the simulation studies to assess the derived relationships did not examine different population values. Only one set of population values were examined in the simulation studies. The derived relationships hold true for the inferred population parameters; they may not hold for sample estimates when the number of clusters is small because of the sampling error. To properly assess the population values, the simulation studies had to utilize a sample generated from a hypothetical population with a very "large" number of clusters. The simulation studies also included the conditions with the smaller number of groups to determine what would constitute the "large" number of clusters.

Another limitation of the simulation is that the accuracy of the results for the estimates is limited to the multilevel mediation model with a balanced data structure. In both simulation studies, I considered a balanced multilevel data, as did Kenny et al. (2003), to mitigate the effect of sampling errors on the REML estimates for the inferred parameters. Note that sampling error causes REML estimates to deviate from the population values. The effect of sampling errors on the population values is more evident when the number of clusters becomes smaller and the multilevel data are unbalanced.

Future Directions

One future direction would to extend the results presented in the current study to a multilevel mediation model with more than three variables. It should be straightforward to extend the CWC2 centering strategy to a 1-1-1 mediation model with multiple *X*s, *M*s, *Y*s.

A second extension of the current study would be to extend the results for the 1-1-1 mediation model to the other types of two-level multilevel mediation models such as 2-2-1, 2-1-1, 2-1-2, 1-2-1, 1-1-2, and 1-2-2. These models are all two level mediation models with a single independent variable, mediator, and response variable.

A third extension of the current study is to extend the results present in this study to multilevel mediation models in which some of *X*'s, *M*'s, and *Y*'s are measured at more than two levels.

A fourth extension of current study would be to use the *observed* cluster means instead of expected values. Using the observed cluster mean could potentially change the derived relationships for the residuals at Level 1; however, it would not be expected to have an effect on the derived relationships for the fixed effects.

A fourth extension would be to derive the covariances between ML estimators in the *uncentered* 1-1-1 mediation model as well as CWC2 1-1-1 mediation models. For example, what is covariance between $\hat{a}$ and $\hat{b}$, $\hat{a}_b$ and $\hat{b}_b$, and $\hat{a}_w$ and $\hat{b}_w$? These covariances can be used to develop CIs to the mediated effects in the multilevel model. Once these covariances are known, then one can obtain CIs for the mediated effects using the distribution of the product method implemented in PRODCLIN program (MacKinnon et al., 2007).

A final future direction would be to investigate different Level-2 covariance structures for a misspecified CWC2 1-1-1 mediation model. Kenny et al. (2003) considered correlations between a_j and b_j. However, the residuals for intercepts and slopes other than a_j and b_j can also be correlated. For example, u_{a_j} and $u_{c_j'}$ can be correlated. Future research needs to address implications of possible correlation between other possible pairs of Level-2 residuals for slopes and intercepts across equations. Note that in a CWC2 1-1-1 mediation model, the Level-2 covariance structure consists of six covariances between intercepts and slopes across equations. In a misspecified CWC2 1-1-1 mediation model, some or all of these covariances can be nonzero. Future research needs to investigate the implications of these nonzero covariances on the assumptions underlying a CWC2 1-1-1 mediation model.

REFERENCES

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.

Alwin, D. F., & Hauser, R. M. (1975). Decomposition of effects in path analysis. *American Sociological Review*, *40*, 37-47.

Anderson, T. W. (2003). *An introduction to multivariate statistical analysis* (3rd ed.). Hoboken, N.J.: Wiley.

Aroian, L. A. (1947). The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics*, *18*(2), 265-271.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182.

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods*, *11*, 142-163. doi:10.1037/1082-989X.11.2.142

Botvin, G. J., & Eng, A. (1980). A comprehensive school-based smoking prevention program. *Journal of School Health*, *50*, 209-213. doi:10.1111/j.1746-1561.1980.tb07378.x

Brook, J. S., Whiteman, M., Finch, S., & Cohen, P. (1995). Aggression, intrapsychic distress, and drug use. *Journal of the American Academy of Child & Adolescent Psychiatry*, *34*, 1076-1084. doi:10.1097/00004583-199508000-00018

Buyse, M., & Molenberghs, G. (1998). Criteria for the validation of surrogate endpoints in randomized experiments. *Biometrics*, *54*, 1014-1029.

Chen, H. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.

Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, H. J., … Long, B. (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American Psychologist*, *48*, 1013-1022. doi:10.1037/0003-066X.48.10.1013

Craig, C. C. (1936). On the frequency function of xy. *The Annals of Mathematical Statistics*, *7*, 1-15. doi:10.1214/aoms/1177732541

Dishion, T. J., & Owen, L. D. (2002). A longitudinal analysis of friendships and substance use: Bidirectional influence from adolescence to adulthood. *Developmental Psychology*, *38*, 480-491. doi:10.1037/0012-1649.38.4.480

Dishion, T. J., & Skaggs, N. M. (2000). An ecological analysis of monthly "bursts" in early adolescent substance use. *Applied Developmental Science*, *4*, 89-97. doi:10.1207/S1532480XADS0402_4

Dooley, D., Prause, J., Ham-Rowbottom, K., & Emptage, N. (2006). Age of alcohol drinking onset: Precursors and the mediation of alcohol disorder. *Journal of Child & Adolescent Substance Abuse*, *15*, 19-37. doi:10.1300/J029v15n02_02

Ellickson, P. L., McCaffrey, D. F., Ghosh-Dastidar, B., & Longshore, D. L. (2003). New inroads in preventing adolescent drug use: Results from a large-scale trial of project ALERT in middle schools. *American Journal of Public Health*, *93*, 1830-1836.

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*, 121-138. doi:10.1037/1082-989X.12.2.121

Epstein, J. A., Bang, H., & Botvin, G. J. (2007). Which psychosocial factors moderate or directly affect substance use among inner-city adolescents? *Addictive Behaviors*, *32*, 700-713. doi:10.1016/j.addbeh.2006.06.011

Guo, Q., Johnson, C. A., Unger, J. B., Lee, L., Xie, B., Chou, C., … Pentz, M. A. (2007). Utility of the theory of reasoned action and theory of planned behavior for predicting Chinese adolescent smoking. *Addictive Behaviors*, *32*, 1066-1081. doi:10.1016/j.addbeh.2006.07.015

Harlow, L., Mulaik, S. A., & Steiger, J. H. (Eds.). (1997). *What if there were no significance tests?* Mahwah, N.J.: Erlbaum.

Hawkins, J. D., Brown, E. C., Oesterle, S., Arthur, M. W., Abbott, R. D., & Catalano, R. F. (2008). Early effects of communities that care on targeted risks and initiation of delinquent behavior and substance use. *Journal of Adolescent Health*, *43*, 15-22. doi:10.1016/j.jadohealth.2008.01.022

Hedeker, D., & Gibbons, R. D. (1994). Random-effects regression models for clustered data with an example from smoking prevention research. *Biometrics*, *50*, 933-944.

Hunt, N. (2000). Pictograms in Excel. *Teaching Statistics*, *22*, 56-58.

Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods*, *8*, 115-128. doi:10.1037/1082-989X.8.2.115

Kreft, I. G. G., de Leeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research*, *30*, 1-21. doi:10.1207/s15327906mbr3001_1

Krull, J. L., & MacKinnon, D. P. (1999). Multilevel mediation modeling in group-based intervention studies. *Evaluation Review*, *23*, 418 -444. doi:10.1177/0193841X9902300404

Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*, *36*, 249-277. doi:10.1207/S15327906MBR3602_06

Lomnicki, Z. A. (1967). On the distribution of products of random variables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *29*, 513-524.

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, *13*, 203-229. doi:10.1037/a0012869

Lynagh, M., Schofield, M. J., & Sanson-Fisher, R. W. (1997). School health promotion programs over the past decade: A review of the smoking, alcohol and solar protection literature. *Health Promotion International*, *12*, 43-60.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York: Erlbaum.

MacKinnon, D. P., Fritz, M. S., Williams, J., & Lockwood, C. M. (2007). Distribution of the product confidence limits for the indirect effect: Program PRODCLIN. *Behavior Research Methods*, *39*, 384-389.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83-104. doi:10.1037//1082-989X.7.1.83

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*, 99-128. doi:10.1207/s15327906mbr3901_4

Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus User's Guide*. Sixth Edition. Los Angeles: Muthén & Muthén.

Neuhaus, J. M., & Kalbfleisch, J. D. (1998). Between- and within-cluster covariate effects in the analysis of clustered data. *Biometrics*, *54*, 638-645.

Orlando, M., Ellickson, P. L., McCaffrey, D. F., & Longshore, D. L. (2005). Mediation analysis of a school-based drug prevention program: Effects of project ALERT. *Prevention Science*, *6*(1), 35-46. doi:10.1007/s11121-005-1251-z

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879-891. doi:10.3758/BRM.40.3.879

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, *15*, 209-233. doi:10.1037/a0020141

Raudenbush, S. W. (1989a). A response to Longford and Plewis. *Multilevel Modeling Newsletter*, *1*, 8-11.

Raudenbush, S. W. (1989b). Centering predictors in multilevel analysis: Choices and consequences. *Multilevel Modeling Newsletter*, *1*, 10-12.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models : Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, *13*, 290-312.

Wilkinson, L., & the Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-604. doi:10.1037/0003-066X.54.8.594

Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, *14*, 301-322. doi:10.1037/a0016972

Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models. *Organizational Research Methods*, *12*, 695 -719. doi:10.1177/1094428108327450

Table 1

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	50	10	0.4725	0.002	0.4	0.049
a	50	10	0.35	-0.001	-0.2	0.024
c'	50	10	0.35	0.004	1.1	0.024
b	50	10	0.35	-0.002	-0.4	0.023
σ_1^2	50	10	2.1225	0.004	0.2	0.083
σ_2^2	50	10	1	-0.002	-0.2	0.005
σ_3^2	50	10	1	0.003	0.3	0.006
$\sigma_{d_{1j}}^2$	50	10	2.1225	-0.032	-1.5	0.361
$\sigma_{d_{2j}}^2$	50	10	1	0.009	0.9	0.053
$\sigma_{d_{3j}}^2$	50	10	1	-0.023	-2.3	0.055
$\sigma_{c_j}^2$	50	10	2.245	-0.032	-1.4	0.433
$\sigma_{a_j}^2$	50	10	1	-0.025	-2.5	0.049
$\sigma_{c'_j}^2$	50	10	1	0.012	1.2	0.069
$\sigma_{b_j}^2$	50	10	1	0.004	0.4	0.052

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 2

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	50	20	0.4725	0.000	0.0	0.049
a	50	20	0.35	0.015	4.2	0.021
c'	50	20	0.35	0.001	0.4	0.022
b	50	20	0.35	-0.002	-0.5	0.022
σ_1^2	50	20	2.1225	0.004	0.2	0.058
σ_2^2	50	20	1	-0.004	-0.4	0.002
σ_3^2	50	20	1	-0.002	-0.2	0.003
$\sigma_{d_{1j}}^2$	50	20	2.1225	0.043	2.0	0.373
$\sigma_{d_{2j}}^2$	50	20	1	0.017	1.7	0.044
$\sigma_{d_{3j}}^2$	50	20	1	0.008	0.8	0.049
$\sigma_{c_j}^2$	50	20	2.245	-0.035	-1.5	0.427
$\sigma_{a_j}^2$	50	20	1	-0.010	-1.0	0.047
$\sigma_{c'_j}^2$	50	20	1	-0.008	-0.8	0.049
$\sigma_{b_j}^2$	50	20	1	0.010	1.0	0.048

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 3

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	50	50	0.4725	-0.009	-1.9	0.045
a	50	50	0.35	0.005	1.4	0.021
c'	50	50	0.35	-0.006	-1.7	0.021
b	50	50	0.35	-0.005	-1.5	0.020
σ_1^2	50	50	2.1225	-0.013	-0.6	0.058
σ_2^2	50	50	1	0.000	0.0	0.001
σ_3^2	50	50	1	0.000	0.0	0.001
$\sigma_{d_{1j}}^2$	50	50	2.1225	-0.037	-1.8	0.339
$\sigma_{d_{2j}}^2$	50	50	1	-0.001	-0.1	0.040
$\sigma_{d_{3j}}^2$	50	50	1	-0.003	-0.3	0.040
$\sigma_{c_j}^2$	50	50	2.245	0.001	0.0	0.456
$\sigma_{a_j}^2$	50	50	1	0.013	1.3	0.044
$\sigma_{c'_j}^2$	50	50	1	0.004	0.4	0.041
$\sigma_{b_j}^2$	50	50	1	-0.010	-1.0	0.044

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 4

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	50	100	0.4725	0.009	1.9	0.047
a	50	100	0.35	-0.004	-1.2	0.020
c'	50	100	0.35	0.008	2.4	0.020
b	50	100	0.35	0.001	0.3	0.021
σ_1^2	50	100	2.1225	0.000	0.0	0.050
σ_2^2	50	100	1	0.001	0.1	0.000
σ_3^2	50	100	1	0.002	0.2	0.000
$\sigma_{d_{1j}}^2$	50	100	2.1225	0.020	0.9	0.328
$\sigma_{d_{2j}}^2$	50	100	1	-0.010	-1.0	0.042
$\sigma_{d_{3j}}^2$	50	100	1	0.017	1.7	0.040
$\sigma_{c_j}^2$	50	100	2.245	0.037	1.7	0.419
$\sigma_{a_j}^2$	50	100	1	0.007	0.7	0.044
$\sigma_{c'_j}^2$	50	100	1	-0.002	-0.2	0.041
$\sigma_{b_j}^2$	50	100	1	-0.003	-0.3	0.042

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 5

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	100	10	0.4725	-0.003	-0.7	0.027
a	100	10	0.35	0.000	0.1	0.013
c'	100	10	0.35	-0.004	-1.2	0.014
b	100	10	0.35	-0.003	-0.8	0.011
σ_1^2	100	10	2.1225	-0.014	-0.7	0.040
σ_2^2	100	10	1	0.000	0.0	0.002
σ_3^2	100	10	1	-0.003	-0.3	0.003
$\sigma_{d_{1j}}^2$	100	10	2.1225	-0.020	-0.9	0.226
$\sigma_{d_{2j}}^2$	100	10	1	0.004	0.4	0.025
$\sigma_{d_{3j}}^2$	100	10	1	0.005	0.5	0.035
$\sigma_{c_j}^2$	100	10	2.245	-0.018	-0.8	0.212
$\sigma_{a_j}^2$	100	10	1	-0.011	-1.1	0.025
$\sigma_{c'_j}^2$	100	10	1	-0.008	-0.8	0.033
$\sigma_{b_j}^2$	100	10	1	-0.007	-0.7	0.028

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 6

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	100	20	0.4725	-0.006	-1.4	0.026
a	100	20	0.35	0.002	0.6	0.009
c'	100	20	0.35	-0.005	-1.5	0.011
b	100	20	0.35	0.003	1.0	0.011
σ_1^2	100	20	2.1225	0.001	0.1	0.029
σ_2^2	100	20	1	0.004	0.4	0.001
σ_3^2	100	20	1	0.002	0.2	0.001
$\sigma_{d_{1j}}^2$	100	20	2.1225	0.028	1.3	0.190
$\sigma_{d_{2j}}^2$	100	20	1	0.002	0.2	0.022
$\sigma_{d_{3j}}^2$	100	20	1	0.014	1.4	0.022
$\sigma_{c_j}^2$	100	20	2.245	-0.013	-0.6	0.229
$\sigma_{a_j}^2$	100	20	1	0.001	0.1	0.024
$\sigma_{c'_j}^2$	100	20	1	-0.001	-0.1	0.025
$\sigma_{b_j}^2$	100	20	1	-0.004	-0.4	0.021

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 7

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Coefficients	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	100	50	0.4725	-0.007	-1.4	0.023
a	100	50	0.35	0.007	1.9	0.011
c'	100	50	0.35	-0.004	-1.1	0.010
b	100	50	0.35	-0.003	-0.8	0.009
σ_1^2	100	50	2.1225	0.001	0.0	0.028
σ_2^2	100	50	1	0.002	0.2	0.000
σ_3^2	100	50	1	0.001	0.1	0.000
$\sigma_{d_{1j}}^2$	100	50	2.1225	-0.022	-1.0	0.179
$\sigma_{d_{2j}}^2$	100	50	1	0.003	0.3	0.024
$\sigma_{d_{3j}}^2$	100	50	1	-0.001	-0.1	0.023
$\sigma_{c_j}^2$	100	50	2.245	0.008	0.4	0.206
$\sigma_{a_j}^2$	100	50	1	0.002	0.2	0.022
$\sigma_{c'_j}^2$	100	50	1	0.008	0.8	0.021
$\sigma_{b_j}^2$	100	50	1	0.000	0.0	0.023

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 8

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	100	100	0.4725	0.006	1.3	0.021
a	100	100	0.35	0.000	-0.1	0.009
c'	100	100	0.35	0.006	1.7	0.009
b	100	100	0.35	0.008	2.2	0.010
σ_1^2	100	100	2.1225	0.004	0.2	0.027
σ_2^2	100	100	1	0.000	0.0	0.000
σ_3^2	100	100	1	-0.001	-0.1	0.000
$\sigma_{d_{1j}}^2$	100	100	2.1225	-0.030	-1.4	0.174
$\sigma_{d_{2j}}^2$	100	100	1	-0.004	-0.4	0.023
$\sigma_{d_{3j}}^2$	100	100	1	-0.009	-0.9	0.021
$\sigma_{c_j}^2$	100	100	2.245	-0.044	-2.0	0.186
$\sigma_{a_j}^2$	100	100	1	-0.008	-0.8	0.021
$\sigma_{c'_j}^2$	100	100	1	-0.004	-0.4	0.023
$\sigma_{b_j}^2$	100	100	1	-0.003	-0.3	0.023

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 9

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	200	10	0.4725	-0.002	-0.5	0.013
a	200	10	0.35	0.010	2.8	0.006
c'	200	10	0.35	-0.005	-1.4	0.007
b	200	10	0.35	-0.008	-2.3	0.006
σ_1^2	200	10	2.1225	-0.004	-0.2	0.021
σ_2^2	200	10	1	0.001	0.1	0.001
σ_3^2	200	10	1	-0.001	-0.1	0.001
$\sigma_{d_{1j}}^2$	200	10	2.1225	0.014	0.6	0.098
$\sigma_{d_{2j}}^2$	200	10	1	0.009	0.9	0.013
$\sigma_{d_{3j}}^2$	200	10	1	-0.001	-0.1	0.016
$\sigma_{c_j}^2$	200	10	2.245	0.028	1.3	0.116
$\sigma_{a_j}^2$	200	10	1	0.007	0.7	0.013
$\sigma_{c'_j}^2$	200	10	1	0.013	1.3	0.016
$\sigma_{b_j}^2$	200	10	1	0.000	0.0	0.011

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 10

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	200	20	0.4725	-0.002	-0.4	0.012
a	200	20	0.35	-0.002	-0.5	0.005
c'	200	20	0.35	0.004	1.1	0.006
b	200	20	0.35	-0.007	-2.1	0.006
σ_1^2	200	20	2.1225	-0.006	-0.3	0.018
σ_2^2	200	20	1	-0.001	-0.1	0.000
σ_3^2	200	20	1	0.001	0.1	0.001
$\sigma_{d_{1j}}^2$	200	20	2.1225	0.017	0.8	0.090
$\sigma_{d_{2j}}^2$	200	20	1	-0.002	-0.2	0.012
$\sigma_{d_{3j}}^2$	200	20	1	0.007	0.7	0.012
$\sigma_{c_j}^2$	200	20	2.245	0.001	0.0	0.109
$\sigma_{a_j}^2$	200	20	1	0.004	0.4	0.012
$\sigma_{c'_j}^2$	200	20	1	0.001	0.1	0.012
$\sigma_{b_j}^2$	200	20	1	0.000	0.0	0.013

Note. The inferred parameters c, $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 11

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	200	50	0.4725	0.002	0.4	0.011
a	200	50	0.35	0.010	2.9	0.005
c'	200	50	0.35	0.000	0.1	0.005
b	200	50	0.35	-0.004	-1.0	0.005
σ_1^2	200	50	2.1225	-0.012	-0.6	0.014
σ_2^2	200	50	1	-0.001	-0.1	0.000
σ_3^2	200	50	1	-0.001	-0.1	0.000
$\sigma_{d_{1j}}^2$	200	50	2.1225	-0.013	-0.6	0.083
$\sigma_{d_{2j}}^2$	200	50	1	-0.009	-0.9	0.010
$\sigma_{d_{3j}}^2$	200	50	1	0.007	0.7	0.011
$\sigma_{c_j}^2$	200	50	2.245	0.023	1.0	0.105
$\sigma_{a_j}^2$	200	50	1	0.005	0.5	0.011
$\sigma_{c'_j}^2$	200	50	1	0.010	1.0	0.011
$\sigma_{b_j}^2$	200	50	1	-0.007	-0.7	0.011

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 12

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	200	100	0.4725	0.001	0.3	0.012
a	200	100	0.35	-0.003	-0.9	0.005
c'	200	100	0.35	0.000	0.0	0.005
b	200	100	0.35	-0.001	-0.3	0.005
σ_1^2	200	100	2.1225	-0.001	-0.1	0.013
σ_2^2	200	100	1	0.001	0.1	0.000
σ_3^2	200	100	1	-0.001	-0.1	0.000
$\sigma_{d_{1j}}^2$	200	100	2.1225	-0.008	-0.4	0.086
$\sigma_{d_{2j}}^2$	200	100	1	0.002	0.2	0.009
$\sigma_{d_{3j}}^2$	200	100	1	0.001	0.1	0.011
$\sigma_{c_j}^2$	200	100	2.245	0.003	0.1	0.104
$\sigma_{a_j}^2$	200	100	1	-0.004	-0.4	0.010
$\sigma_{c'_j}^2$	200	100	1	0.003	0.3	0.011
$\sigma_{b_j}^2$	200	100	1	-0.002	-0.2	0.011

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 13

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	500	10	0.4725	0.005	0.1	0.006
a	500	10	0.35	-0.001	0.4	0.002
c'	500	10	0.35	0.003	1.0	0.003
b	500	10	0.35	0.002	0.7	0.002
σ_1^2	500	10	2.1225	-0.002	0.1	0.008
σ_2^2	500	10	1	0.000	0.1	0.001
σ_3^2	500	10	1	0.001	0.1	0.001
$\sigma_{d_{1j}}^2$	500	10	2.1225	-0.013	-0.6	0.038
$\sigma_{d_{2j}}^2$	500	10	1	0.002	0.2	0.005
$\sigma_{d_{3j}}^2$	500	10	1	0.002	0.2	0.006
$\sigma_{c_j}^2$	500	10	2.245	-0.003	0.2	0.044
$\sigma_{a_j}^2$	500	10	1	-0.002	0.2	0.005
$\sigma_{c'_j}^2$	500	10	1	0.004	0.4	0.007
$\sigma_{b_j}^2$	500	10	1	-0.007	-1.7	0.005

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 14

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	500	20	0.4725	0.002	0.4	0.005
a	500	20	0.35	-0.001	0.3	0.002
c'	500	20	0.35	-0.002	-0.6	0.002
b	500	20	0.35	-0.001	0.2	0.002
σ_1^2	500	20	2.1225	0.002	0.1	0.006
σ_2^2	500	20	1	0.000	0	0.000
σ_3^2	500	20	1	0.001	0.1	0.000
$\sigma_{d_{1j}}^2$	500	20	2.1225	-0.009	-0.4	0.040
$\sigma_{d_{2j}}^2$	500	20	1	0.003	0.3	0.004
$\sigma_{d_{3j}}^2$	500	20	1	-0.002	0.2	0.005
$\sigma_{c_j}^2$	500	20	2.245	-0.019	-0.8	0.041
$\sigma_{a_j}^2$	500	20	1	-0.007	-0.7	0.005
$\sigma_{c'_j}^2$	500	20	1	-0.007	-0.7	0.005
$\sigma_{b_j}^2$	500	20	1	-0.001	0.1	0.004

Note. The inferred parameters c, $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 15

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	500	50	0.4725	0.002	0.4	0.005
a	500	50	0.35	0.001	0.2	0.002
c'	500	50	0.35	0.006	1.6	0.002
b	500	50	0.35	-0.004	-1.3	0.002
σ_1^2	500	50	2.1225	-0.004	0.2	0.004
σ_2^2	500	50	1	0.001	0.1	0.000
σ_3^2	500	50	1	0.001	0.1	0.000
$\sigma_{d_{1j}}^2$	500	50	2.1225	0.027	1.3	0.041
$\sigma_{d_{2j}}^2$	500	50	1	0.002	0.2	0.004
$\sigma_{d_{3j}}^2$	500	50	1	0.007	0.7	0.005
$\sigma_{c_j}^2$	500	50	2.245	-0.014	-0.6	0.034
$\sigma_{a_j}^2$	500	50	1	0.004	0.4	0.004
$\sigma_{c'_j}^2$	500	50	1	-0.009	-0.9	0.005
$\sigma_{b_j}^2$	500	50	1	-0.002	0.2	0.004

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 16

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	500	100	0.4725	-0.001	0.3	0.004
a	500	100	0.35	0.004	1.1	0.002
c'	500	100	0.35	-0.002	-0.6	0.002
b	500	100	0.35	-0.002	0.5	0.002
σ_1^2	500	100	2.1225	-0.003	0.2	0.005
σ_2^2	500	100	1	-0.001	0.1	0.000
σ_3^2	500	100	1	0.000	0	0.000
$\sigma_{d_{1j}}^2$	500	100	2.1225	0.007	0.3	0.033
$\sigma_{d_{2j}}^2$	500	100	1	-0.005	-0.5	0.005
$\sigma_{d_{3j}}^2$	500	100	1	0.010	0.1	0.006
$\sigma_{c_j}^2$	500	100	2.245	0.020	0.9	0.034
$\sigma_{a_j}^2$	500	100	1	0.012	1.2	0.003
$\sigma_{c'_j}^2$	500	100	1	0.004	0.4	0.004
$\sigma_{b_j}^2$	500	100	1	-0.001	-0.1	0.003

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 17

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	1000	10	0.4725	0.000	0.0	0.003
a	1000	10	0.35	0.000	0.0	0.001
c'	1000	10	0.35	0.003	0.7	0.001
b	1000	10	0.35	0.002	0.5	0.001
σ_1^2	1000	10	2.1225	0.009	0.4	0.004
σ_2^2	1000	10	1	0.000	0.0	0.000
σ_3^2	1000	10	1	0.000	0.0	0.000
$\sigma_{d_{1j}}^2$	1000	10	2.1225	0.012	0.6	0.016
$\sigma_{d_{2j}}^2$	1000	10	1	-0.002	-0.2	0.002
$\sigma_{d_{3j}}^2$	1000	10	1	0.004	0.4	0.003
$\sigma_{c_j}^2$	1000	10	2.245	-0.004	-0.2	0.019
$\sigma_{a_j}^2$	1000	10	1	0.004	0.4	0.002
$\sigma_{c'_j}^2$	1000	10	1	-0.008	-0.8	0.004
$\sigma_{b_j}^2$	1000	10	1	0.004	0.4	0.003

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 18

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	1000	20	0.4725	0.001	0.3	0.002
a	1000	20	0.35	0.003	0.9	0.001
c'	1000	20	0.35	0.000	-0.1	0.001
b	1000	20	0.35	0.003	0.9	0.001
σ_1^2	1000	20	2.1225	0.007	0.3	0.003
σ_2^2	1000	20	1	0.000	0.0	0.000
σ_3^2	1000	20	1	0.001	0.1	0.000
$\sigma_{d_{1j}}^2$	1000	20	2.1225	0.017	0.8	0.016
$\sigma_{d_{2j}}^2$	1000	20	1	0.000	0.0	0.003
$\sigma_{d_{3j}}^2$	1000	20	1	0.005	0.5	0.002
$\sigma_{c_j}^2$	1000	20	2.245	0.009	0.4	0.020
$\sigma_{a_j}^2$	1000	20	1	0.001	0.1	0.002
$\sigma_{c'_j}^2$	1000	20	1	-0.001	-0.1	0.003
$\sigma_{b_j}^2$	1000	20	1	0.005	0.5	0.002

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 19

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	1000	50	0.4725	-0.002	-0.5	0.003
a	1000	50	0.35	-0.005	-1.3	0.001
c'	1000	50	0.35	-0.002	-0.5	0.001
b	1000	50	0.35	-0.002	-0.5	0.001
σ_1^2	1000	50	2.1225	-0.004	-0.2	0.003
σ_2^2	1000	50	1	0.001	0.1	0.000
σ_3^2	1000	50	1	0.000	0.0	0.000
$\sigma_{d_{1j}}^2$	1000	50	2.1225	0.006	0.3	0.018
$\sigma_{d_{2j}}^2$	1000	50	1	-0.001	-0.1	0.002
$\sigma_{d_{3j}}^2$	1000	50	1	0.001	0.1	0.002
$\sigma_{c_j}^2$	1000	50	2.245	-0.003	-0.1	0.019
$\sigma_{a_j}^2$	1000	50	1	0.000	0.0	0.002
$\sigma_{c'_j}^2$	1000	50	1	-0.002	-0.2	0.002
$\sigma_{b_j}^2$	1000	50	1	-0.002	-0.2	0.002

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 20

Bias, Percentage Bias, and MSE for Population Parameters in Uncentered 1-1-1 Mediation Model

Parameters	Number of groups	Group size	Population values	Bias	Percentage bias	*MSE*
c	1000	100	0.4725	-0.003	-0.6	0.002
a	1000	100	0.35	0.001	0.3	0.001
c'	1000	100	0.35	-0.005	-1.4	0.001
b	1000	100	0.35	-0.002	-0.6	0.001
σ_1^2	1000	100	2.1225	0.001	0.1	0.002
σ_2^2	1000	100	1	0.000	0.0	0.000
σ_3^2	1000	100	1	0.000	0.0	0.000
$\sigma_{d_{1j}}^2$	1000	100	2.1225	0.025	1.2	0.016
$\sigma_{d_{2j}}^2$	1000	100	1	0.004	0.4	0.002
$\sigma_{d_{3j}}^2$	1000	100	1	0.005	0.5	0.002
$\sigma_{c_j}^2$	1000	100	2.245	0.018	0.8	0.018
$\sigma_{a_j}^2$	1000	100	1	-0.002	-0.2	0.002
$\sigma_{c'_j}^2$	1000	100	1	0.001	0.1	0.002
$\sigma_{b_j}^2$	1000	100	1	0.003	0.3	0.002

Note. The inferred parameters c , $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, and σ_1^2 (shown in gray rows) are not free. They are a function of other parameters.

Table 21

Bias, Percentage Bias, and MSE for Population Parameters in CWC2 1-1-1 Mediation Model (Group Size=10)

Parameters	Number of groups	Population values	Bias	Percentage bias	*MSE*
d_1	50	0.575	0.012	2.0	0.027
d_2	50	0.500	0.017	3.4	0.020
d_3	50	0.500	-0.006	-1.1	0.032
c_w	50	0.473	-0.006	-1.2	0.052
c_b	50	0.173	0.019	10.8	0.031
a_w	50	0.350	0.002	0.5	0.021
a_b	50	0.150	-0.008	-5.5	0.025
c'_w	50	0.350	-0.002	-0.7	0.028
c'_b	50	0.150	0.009	6.0	0.031
b_w	50	0.350	0.005	1.4	0.020
b_b	50	0.150	0.021	14.0	0.018
σ_1^2	50	2.123	-0.011	-0.5	0.086
σ_2^2	50	1.000	-0.003	-0.3	0.005
σ_3^2	50	1.000	-0.008	-0.8	0.006
$\sigma_{d_{1j}}^2$	50	1.023	-0.033	-3.2	0.056
$\sigma_{d_{2j}}^2$	50	1.000	0.012	1.2	0.050
$\sigma_{d_{3j}}^2$	50	1.000	-0.008	-0.8	0.050
$\sigma_{c_j}^2$	50	2.315	-0.056	-2.4	0.386
$\sigma_{a_j}^2$	50	1.000	-0.008	-0.8	0.049
$\sigma_{c'_j}^2$	50	1.000	0.020	2.0	0.061
$\sigma_{b_j}^2$	50	1.000	-0.015	-1.5	0.048
$\sigma_{c'_j, b_j}$	50	0.100	0.012	11.6	0.027
σ_{d_{2j}, a_j}	50	0.100	-0.017	-16.6	0.029
σ_{d_{3j}, b_j}	50	0.100	-0.002	-1.8	0.022
σ_{d_{3j}, c'_j}	50	0.100	0.009	9.1	0.033
σ_{d_{1j}, c_j}	50	0.140	0.016	11.3	0.080

Note. CWC2= Centering within cluster with the cluster mean as a Level-2 predictor of intercept. The inferred parameters d_1, c_w, c_b, $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, σ_1^2, and σ_{d_{1j},c_j} (shown in gray rows) are not free. These parameters are a function of other parameters.

Table 22

Bias, Percentage Bias, and MSE for Population Parameters in CWC2 1-1-1 Mediation Model (Group Size=10)

Parameters	Number of groups	Population values	Bias	Percentage bias	*MSE*
d_1	100	0.575	0.000	0.1	0.013
d_2	100	0.500	-0.007	-1.5	0.010
d_3	100	0.500	0.005	1.0	0.014
c_w	100	0.473	-0.019	-4.0	0.028
c_b	100	0.173	-0.007	-4.1	0.013
a_w	100	0.350	-0.008	-2.3	0.009
a_b	100	0.150	-0.015	-9.9	0.010
c'_w	100	0.350	-0.005	-1.4	0.014
c'_b	100	0.150	-0.005	-3.5	0.013
b_w	100	0.350	-0.013	-3.6	0.013
b_b	100	0.150	-0.002	-1.1	0.008
σ_1^2	100	2.123	-0.010	-0.5	0.041
σ_2^2	100	1.000	-0.001	-0.1	0.002
σ_3^2	100	1.000	-0.005	-0.5	0.003
$\sigma_{d_{1j}}^2$	100	1.023	-0.014	-1.3	0.038
$\sigma_{d_{2j}}^2$	100	1.000	0.010	1.0	0.024
$\sigma_{d_{3j}}^2$	100	1.000	0.018	1.8	0.029
$\sigma_{c_j}^2$	100	2.315	-0.055	-2.4	0.245
$\sigma_{a_j}^2$	100	1.000	0.005	0.5	0.026
$\sigma_{c'_j}^2$	100	1.000	0.027	2.7	0.036
$\sigma_{b_j}^2$	100	1.000	-0.005	-0.5	0.026
$\sigma_{c'_j, b_j}$	100	0.100	0.014	14.3	0.013
σ_{d_{2j}, a_j}	100	0.100	0.009	9.4	0.011
σ_{d_{3j}, b_j}	100	0.100	0.007	7.4	0.014
σ_{d_{3j}, c'_j}	100	0.100	-0.006	-5.9	0.015
σ_{d_{1j}, c_j}	100	0.140	-0.002	-1.6	0.040

Note. CWC2= Centering within cluster with the cluster mean as a Level-2 predictor of intercept. The inferred parameters d_1, c_w, c_b, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, σ^2_1, and σ_{d_{1j},c_j} (shown in gray rows) are not free. These parameters are a function of other parameters.

Table 23

Bias, Percentage Bias, and MSE for Population Parameters in CWC2 1-1-1 Mediation Model (Group Size=10)

Parameters	Number of groups	Population values	Bias	Percentage bias	*MSE*
d_1	200	0.575	0.001	0.3	0.006
d_2	200	0.500	0.007	1.5	0.006
d_3	200	0.500	0.008	1.5	0.007
c_w	200	0.473	-0.001	-0.2	0.013
c_b	200	0.173	0.001	0.6	0.006
a_w	200	0.350	-0.006	-1.7	0.007
a_b	200	0.150	0.006	4.0	0.005
c'_w	200	0.350	-0.002	-0.5	0.007
c'_b	200	0.150	0.002	1.4	0.006
b_w	200	0.350	-0.002	-0.6	0.005
b_b	200	0.150	-0.011	-7.5	0.004
σ_1^2	200	2.123	0.008	0.4	0.021
σ_2^2	200	1.000	0.004	0.4	0.001
σ_3^2	200	1.000	0.001	0.1	0.001
$\sigma_{d_{1j}}^2$	200	1.023	-0.052	-5.1	0.015
$\sigma_{d_{2j}}^2$	200	1.000	0.001	0.1	0.013
$\sigma_{d_{3j}}^2$	200	1.000	-0.013	-1.3	0.012
$\sigma_{c_j}^2$	200	2.315	-0.093	-4.0	0.097
$\sigma_{a_j}^2$	200	1.000	-0.001	-0.1	0.014
$\sigma_{c'_j}^2$	200	1.000	0.002	0.2	0.017
$\sigma_{b_j}^2$	200	1.000	-0.007	-0.7	0.012
$\sigma_{c'_j, b_j}$	200	0.100	0.006	5.6	0.007
σ_{d_{2j}, a_j}	200	0.100	-0.012	-1.2	0.007
σ_{d_{3j}, b_j}	200	0.100	0.005	4.9	0.007
σ_{d_{3j}, c'_j}	200	0.100	-0.004	-4.0	0.008
σ_{d_{1j}, c_j}	200	0.140	-0.015	-10.5	0.020

Note. CWC2= Centering within cluster with the cluster mean as a Level-2 predictor of intercept. The inferred parameters d_1, c_w, c_b, $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, σ_1^2, and σ_{d_{1j},c_j} (shown in gray rows) are not free. These parameters are a function of other parameters.

Table 24

Bias, Percentage Bias, and MSE for Population Parameters in CWC2 1-1-1 Mediation Model (Group Size=10)

Parameters	Number of groups	Population values	Bias	Percentage bias	*MSE*
d_1	500	0.575	-0.001	-0.1	0.002
d_2	500	0.500	-0.007	-1.5	0.003
d_3	500	0.500	-0.002	-0.4	0.002
c_w	500	0.473	-0.007	-1.5	0.005
c_b	500	0.173	0.000	0	0.003
a_w	500	0.350	-0.008	-2.4	0.003
a_b	500	0.150	-0.001	-0.3	0.002
c'_w	500	0.350	0.003	0.8	0.002
c'_b	500	0.150	0.000	0.3	0.003
b_w	500	0.350	0.003	0.9	0.002
b_b	500	0.150	0.005	3.2	0.002
σ_1^2	500	2.123	0.004	0.2	0.009
σ_2^2	500	1.000	0.000	0	0.001
σ_3^2	500	1.000	-0.002	-0.2	0.001
$\sigma_{d_{1j}}^2$	500	1.023	-0.026	-2.5	0.007
$\sigma_{d_{2j}}^2$	500	1.000	-0.003	-0.3	0.005
$\sigma_{d_{3j}}^2$	500	1.000	0.004	0.4	0.005
$\sigma_{c_j}^2$	500	2.315	-0.066	-2.9	0.049
$\sigma_{a_j}^2$	500	1.000	0.000	0	0.005
$\sigma_{c'_j}^2$	500	1.000	-0.002	-0.2	0.007
$\sigma_{b_j}^2$	500	1.000	-0.002	-0.2	0.005
$\sigma_{c'_j, b_j}$	500	0.100	-0.004	-3.8	0.003
σ_{d_{2j}, a_j}	500	0.100	-0.002	-1.7	0.003
σ_{d_{3j}, b_j}	500	0.100	0.000	0.1	0.003
σ_{d_{3j}, c'_j}	500	0.100	0.003	3.3	0.003
σ_{d_{1j}, c_j}	500	0.140	0.013	9.1	0.010

Note. CWC2= Centering within cluster with the cluster mean as a Level-2 predictor of intercept. The inferred parameters d_1, c_w, c_b, $\sigma^2_{d_{1j}}$, $\sigma^2_{c_j}$, σ^2_1, and σ_{d_{1j},c_j} (shown in gray rows) are not free. These parameters are a function of other parameters.

Table 25

Bias, Percentage Bias, and MSE for Population Parameters in CWC2 1-1-1 Mediation Model (Group Size=10)

Parameters	Number of groups	Population values	Bias	Percentage bias	*MSE*
d_1	1000	0.575	0.003	0.6	0.001
d_2	1000	0.500	0.006	1.2	0.001
d_3	1000	0.500	0.001	0.2	0.002
c_w	1000	0.473	0.003	0.7	0.003
c_b	1000	0.173	0.002	1.2	0.001
a_w	1000	0.350	0.004	1.1	0.001
a_b	1000	0.150	-0.004	-2.9	0.001
c'_w	1000	0.350	-0.001	-0.4	0.001
c'_b	1000	0.150	0.001	1.0	0.001
b_w	1000	0.350	0.004	1.2	0.001
b_b	1000	0.150	0.002	1.5	0.001
σ_1^2	1000	2.123	0.016	0.7	0.004
σ_2^2	1000	1.000	0.000	0	0.000
σ_3^2	1000	1.000	-0.001	-0.1	0.000
$\sigma_{d_{1j}}^2$	1000	1.023	-0.028	-2.7	0.003
$\sigma_{d_{2j}}^2$	1000	1.000	0.002	0.2	0.002
$\sigma_{d_{3j}}^2$	1000	1.000	0.001	0.1	0.002
$\sigma_{c_j}^2$	1000	2.315	-0.061	-2.6	0.016
$\sigma_{a_j}^2$	1000	1.000	0.001	0.1	0.003
$\sigma_{c'_j}^2$	1000	1.000	0.006	0.6	0.003
$\sigma_{b_j}^2$	1000	1.000	0.003	0.3	0.002
$\sigma_{c'_j, b_j}$	1000	0.100	0.003	3.3	0.001
σ_{d_{2j}, a_j}	1000	0.100	0.001	0.9	0.001
σ_{d_{3j}, b_j}	1000	0.100	0.006	5.8	0.001
σ_{d_{3j}, c'_j}	1000	0.100	0.001	1.1	0.001
σ_{d_{1j}, c_j}	1000	0.140	0.002	1.7	0.004

Note. CWC2= Centering within cluster with the cluster mean as a Level-2 predictor of intercept. The inferred parameters d_1, c_w, c_b, $\sigma_{d_{1j}}^2$, $\sigma_{c_j}^2$, σ_1^2, and σ_{d_{1j},c_j} (shown in gray rows) are not free. These parameters are a function of other parameters.

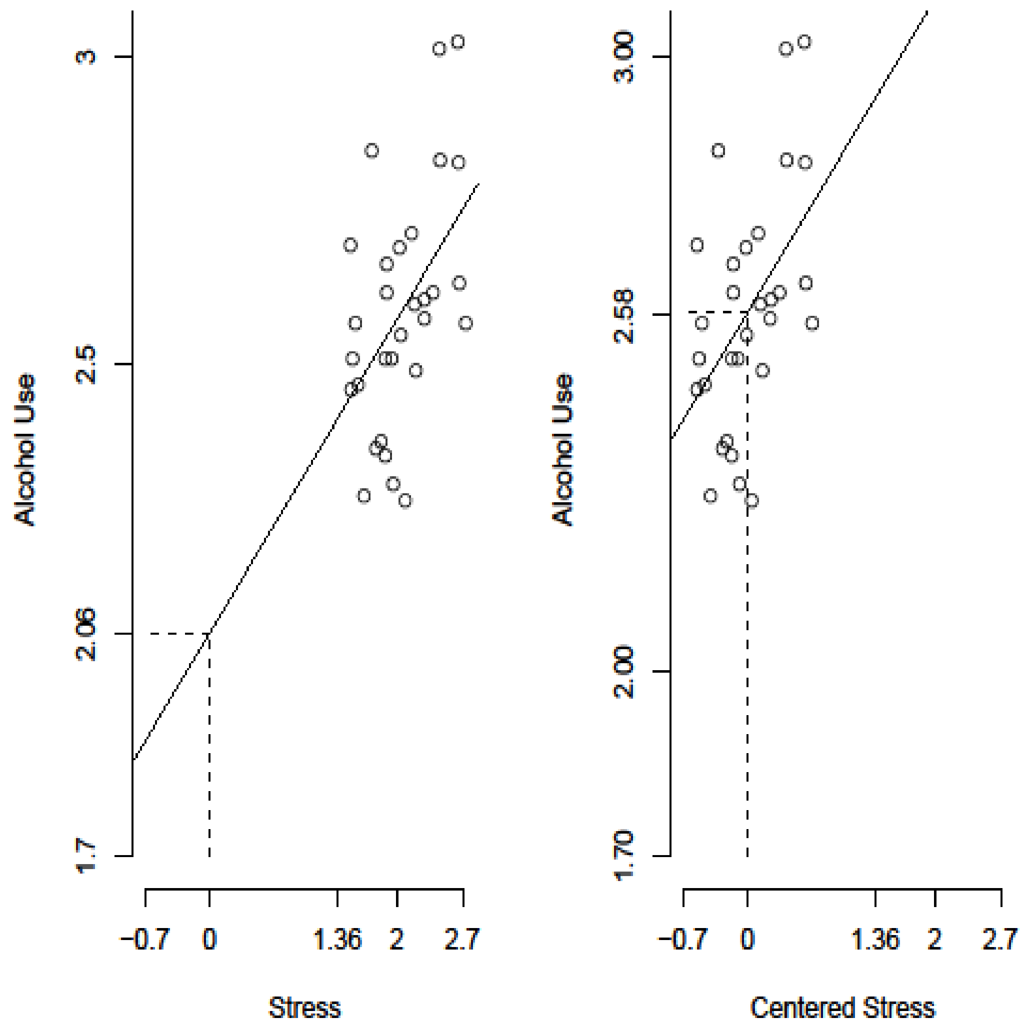

Figure 1. Within-person regression lines for the relationship between daily stress and daily alcohol use for Person 2. The graph on the left shows the regression line for uncentered daily stress and the graph on the right shows the regression line for centered daily stress.

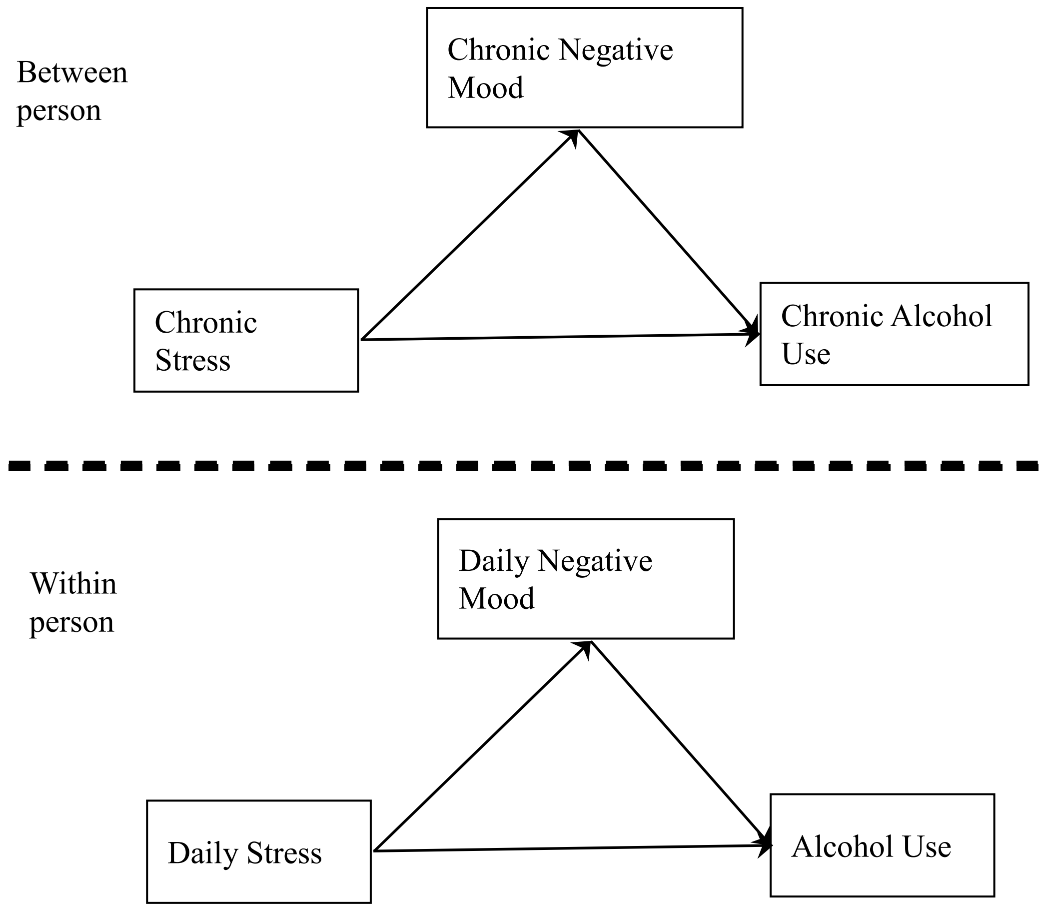

Figure 2. Separation of between-person and within-person mediated effects in the CWC2 multilevel mediation model.

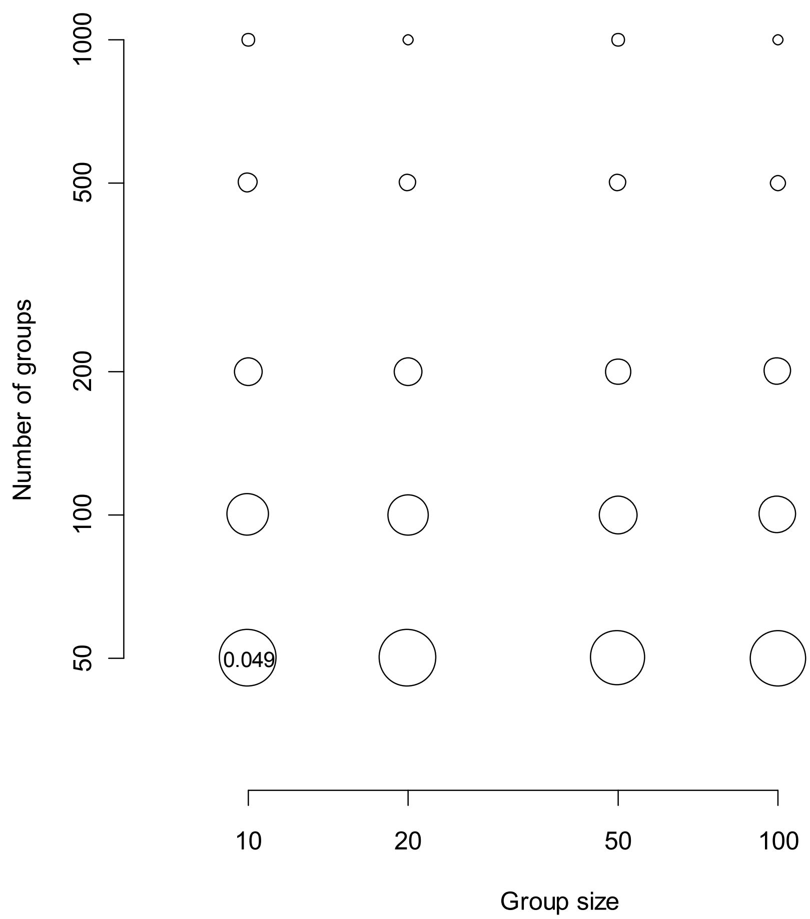

Figure 3. *MSE*s for the REML estimator of parameter c. The area of each bubble is proportional to the magnitude of the associated *MSE*. Larger bubbles indicate larger *MSE*s.

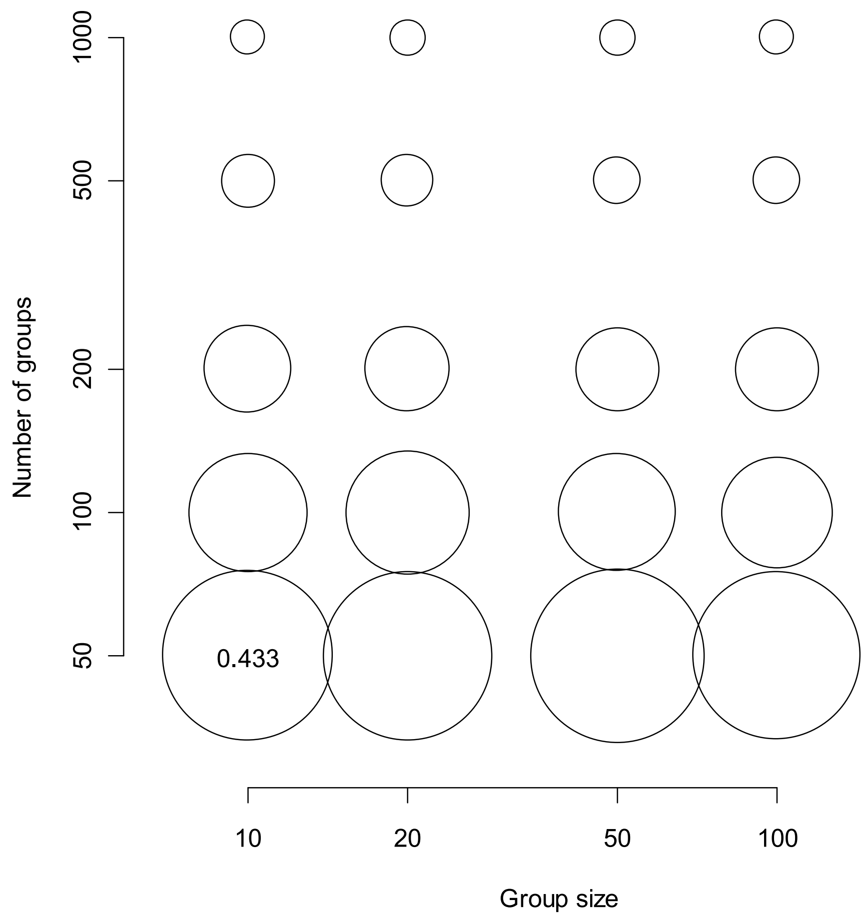

*Figure 4. MSE*s for the REML estimator of parameter σ_c^2. The area of each bubble is proportional to the magnitude of the associated *MSE*. Larger bubbles indicate larger *MSE*s.

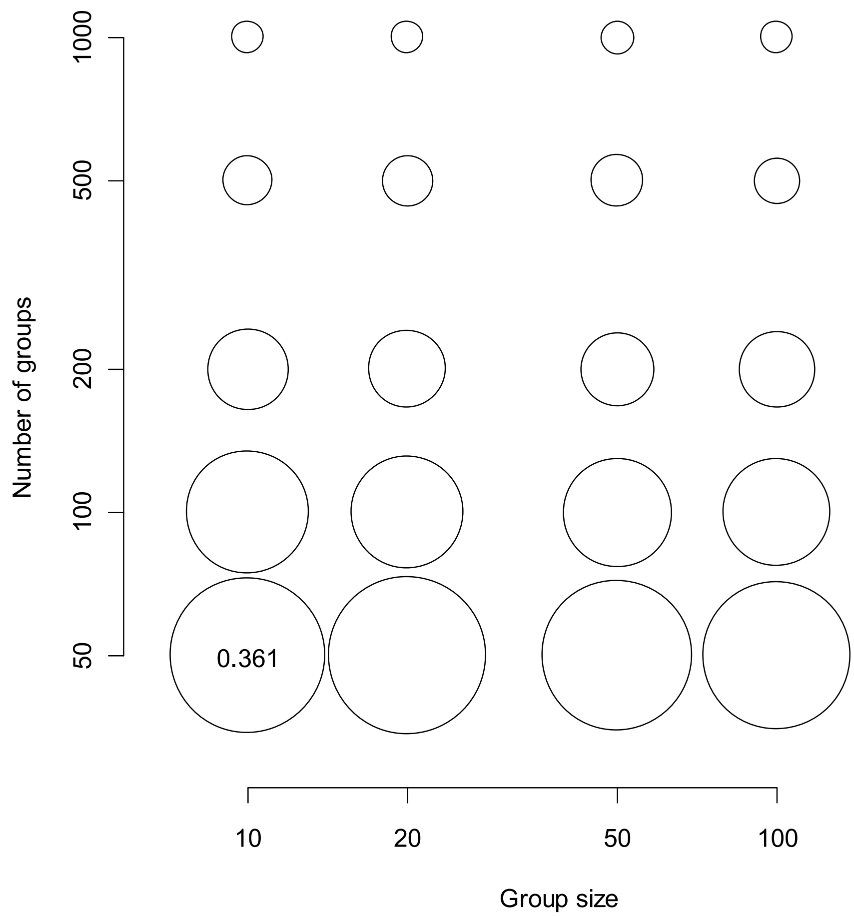

*Figure 5. MSE*s for the REML estimator of parameter $\sigma^2_{d_{1j}}$. The area of each bubble is proportional to the magnitude of the associated *MSE*. Larger bubbles indicate larger *MSE*s.

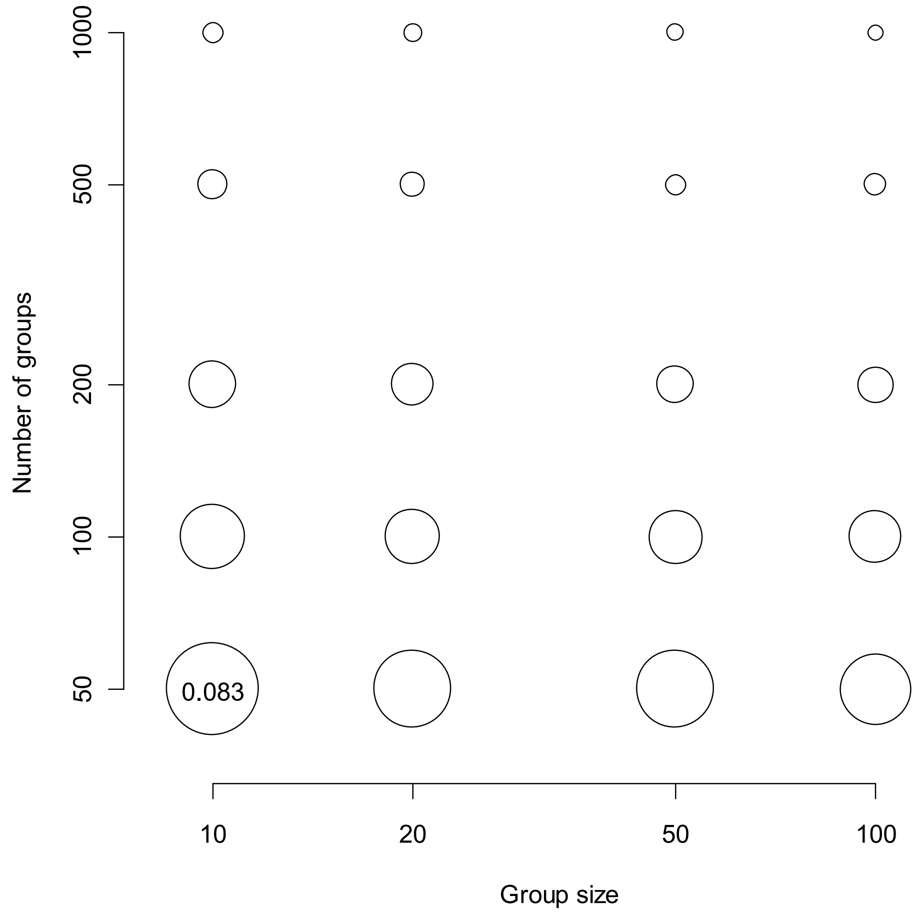

*Figure 6. MSE*s for the REML estimator of parameter σ_1^2. The area of each bubble is proportional to the magnitude of the associated *MSE*. Larger bubbles indicate larger *MSE*s.

111

This document was generated using the Graduate College Format Advising tool. Please turn a copy of this page in when you submit your document to Graduate College format advising. You may discard this page once you have printed your final document. DO NOT TURN THIS PAGE IN WITH YOUR FINAL DOCUMENT!

111

111

111